Validity and Reliability of HLD Treatment Need Indices Currently Used in the United States

Timothee Cousin

A thesis submitted

in partial fulfillment

of the requirements for the degree of

Master of Science in Dentistry

University of Washington

2020

Committee:

Burcu Bayirli

Geoff Greenlee

Gerald Nelson

Program Authorized to Offer Degree:

Department of Orthodontics

University of Washington

Abstract

Validity and Reliability of HLD Treatment Need Indices Currently Used in the United States

Timothee Cousin

Chair of the Supervisory Committee:

Burcu Bayirli

Department of Orthodontics

Introduction: State Medicaid agencies rely on treatment need indices to identify patients with orthodontic medical necessity. The Handicapping Labiolingual Deviations (HLD) index is the most common orthodontic index in the United States. States which have adopted the HLD index have modified it without supporting validation studies. The purpose of this study was to validate all state-modified HLD indices currently used in the United States and to investigate the implementation of the Washington state HLD modification, HLD(WA). **Methods:** To validate all current HLD indices, a panel of 17 orthodontists was recruited to assess the malocclusion of 204 consecutive patient records obtained from the University of Washington Department of Orthodontics archives. State HLD indices were subsequently applied to all 204 patients to predict state-specific orthodontic treatment need decisions. ROC curves were generated for each state-modified HLD indices to determine each index' diagnostic value. We also computed the optimal point thresholds for all state-modified HLD indices. To evaluate the implementation of HLD(WA), the HLD prior-authorizations submitted on behalf of 418 Medicaid-eligible patients by five orthodontic clinics were gathered and compared to corresponding Washington state Medicaid program treatment decisions. We evaluated approval and denial rates among different clinics and compared applications approved and denied according to index guidelines versus reviewer override. **Results:** There were many qualitative differences among state-modified HLD indices. State-modified HLD indices had a mean between-state agreement of 79.1%, a mean sensitivity of 41.6%, and a mean specificity of 92.9%. The optimal point thresholds were lower than the current point thresholds for all state-modified HLD indices. When the implementation of HLD(WA) was evaluated, we found a mean approval rate of 35.4% and a mean denial rate of 52.4%. While total mean approval and denial rates were not significantly different among clinic types, there were significant differences among clinic types in the number of applications approved and denied according to index guideline versus reviewer override ($p<.01$). **Conclusions:** Our results suggest that state-modified HLD indices fail to identify a considerable percentage of handicapping malocclusions. There is also a lack of consensus among state Medicaid programs on what constitutes a handicapping malocclusion. In addition, our study indicates that there are differences in approval and denial rates among clinic types. These findings may exacerbate access to orthodontic care disparities among Medicaid-eligible populations.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank several individuals who provided valuable insights in the preparatory stages and execution of this study. Charles Spiekerman helped design the study from a biostatistical perspective. Michael Nash executed a large portion of the biostatistical analysis in this work. His diligence, responsiveness, and enthusiasm were invaluable. I would also like to thank Drs. Christopher Riolo, Greg Huang, and David Turpin, who in their consulting roles, provided thought-provoking insights, feedback, and advice in all stages of this study. Finally, I would like to thank the University of Washington Orthodontic Alumni Association (UWOAA) and the UW School of Dentistry for their financial support.

BACKGROUND

Introduction

Five types of occlusal indices are currently in use in orthodontic practice: diagnostic, epidemiologic, orthodontic treatment need, orthodontic treatment outcome, and orthodontic treatment complexity indices.[1] Orthodontic treatment need indices are used to prioritize orthodontic care where there is limited funds available to treat a population seeking orthodontic treatment, as is often the case in public health programs. Because malocclusion is a deviation from normal or defined "ideal" anatomy rather than a disease state, it is often treated with great subjectivity. Orthodontic treatment need indices were therefore developed to minimize the subjectivity inherent in diagnosing and treating malocclusions and to provide a more objective means of quantifying malocclusion severity and prioritizing orthodontic treatment care.

Identification of handicapping malocclusions has long been a challenge for orthodontists and public health officials alike. Since their inception in the 1950s, orthodontic treatment need indices have gradually gained in popularity among government health agencies and insurance providers. Today, they are especially common among Northern European countries where orthodontic treatment is heavily subsidized by public health programs, and a means of prioritizing treatment objectively was perceived as necessary for the equitable distribution of orthodontic care.[1] In the United States, the incentive for the development of orthodontic need indices was spurred by the introduction of Medicaid statutes in the early 1960s directing that each state make orthodontic treatment accessible to patients with medically necessary handicapping malocclusions.[2,3] A series of treatment need indices such as Draker's Handicapping Labiolingual Deviation (HLD) index,[4,5] Summer's Occlusal Index,[6] Grainger's Treatment Priority Index,[7] and Salzmann's Handicapping Malocclusion Assessment Record (HMAR)[8] were developed to give dental providers the necessary administrative tools to identify and treat handicapping malocclusions using public funds. In 1969, the American Association of Orthodontists adopted the Salzmann index as its objective qualifier of choice.[9]

In the decades that followed, many more indices were developed as researchers and clinicians sought to simplify the index component measurements and maximize index validity and inter-examiner reproducibility. As such, the Dental Aesthetic Index (DAI),[10] the Index of Orthodontic Treatment Need (IOTN),[11] and the Index of Complexity, Outcome, and Need (ICON)[12] were developed. While European countries continued to expand the use of orthodontic treatment need indices, these efforts were hampered in the United States due, in part, to a position taken by the American Association of Orthodontists (AAO) in a 1990 bulletin rescinding its earlier adoption of the Salzmann index and declaring that it "does not support any index, rating classification or coding system as scientifically valid measures of the need for orthodontic treatment".[13] Nearly a decade elapsed before new research was published on the topic of orthodontic treatment need indices and Parker began his work on modifying Draker's original HLD.[14] At this time, several lawsuits had been filed against individual states, including the state of California, for failing to comply with the orthodontic provisions of the Medicaid statutes. In 1989, Parker was sought to act as expert witness in a lawsuit against the state of California. In the process, Parker oversaw the addition of several modifications to Draker's HLD index used by the state of California, as part of the lawsuit settlement including overjet in excess of 9mm, deep impinging bites that are destroying the soft tissue of the palate, and crossbites of individual anterior teeth when destruction of soft tissue is present.[14] After a second lawsuit against the state of California, two additional component measurements were added as part of the settlement: unilateral posterior crossbite and reverse overjet greater than 3.5mm.[14] Thus, the California modification HLD, or HLD(CalMod), was developed and subsequently validated.[15–17]

At the center of the issue of orthodontic treatment need, is the lack of a clearly defined set of orthodontic diagnoses that could be universally accepted as constituting a "medically-necessary handicapping malocclusion" as required by the Medicaid statutes. In an effort to help clinicians and public health officials come to a consensus as to which types of orthodontic diagnoses should receive public funding for orthodontic treatment, the AAO released a statement in July of 2015 defining medically necessary orthodontic care as *"the treatment of a malocclusion (including craniofacial abnormalities/anomalies) that compromise the patient's physical, emotional or dental health. This treatment should be based on a comprehensive assessment and diagnosis done by an orthodontist, in consultation with other health care providers when indicated"*.[18] In this same bulletin, the AAO proposed its own criteria consisting of a set of diagnostic auto-qualifiers, many of which were already part of the HLD (CalMod) developed by Parker.[18] In 2017, the AAO updated its list of proposed auto-qualifiers, informed its members of its intention to qualify 10% of all cases as medically necessary, and reported that a retrospective study was underway to achieve this goal.[19] Despite advancing its position on medical necessity, the AAO's statements raise more questions than answers. For example, if medical necessity includes *"malocclusions which compromise the patient's physical, emotional, and dental health*," then it would follow that a widespread orthodontic need index should include characteristics of malocclusion as well as psychological, periodontal , and other pertinent components. In addition, it is unclear if there is sufficient evidence to support the claim that an index should qualify 10% of patients as having medical necessity. Data from NHANES III indicate that 10.2-21.5% of American children and youth had definite orthodontic need as measured by the IOTN.[20] The AAO's proposed-index would potentially leave out a substantial number of patients predicted to have medically-necessary orthodontic need. While the AAO's adoption of a universal orthodontic treatment need index would be a step forward towards increasing access to care for orthodontic patients in the United States, this process should be developed on a foundation of strong scientific evidence rather than political debate.

Validation and Reliability Studies

Currently, two major orthodontic treatment need indices are in use in the United States: the Salzmann and the HLD indices. Together, the Salzmann index, HLD index, and their modifications were used by 27 states as of 2015.[9] While the lawsuit driven modification process that took place in the 1990s has helped standardize the HLD index from state to state, careful evaluation of individual state indices listed as using the HLD index reveals variations in the component measurements as well as the cut-off score for eligibility.

Even though research has been conducted to validate indices like the IOTN prior to their implementation by European public health agencies, there is a paucity of validation studies concerning those indices used by Medicaid agencies and insurance providers in the United States, including the HLD index and its many modifications. To ensure that orthodontic need indices reliably and accurately identify patients with great orthodontic need, validation and reliability studies should be conducted prior to index implementation.

Reliability is the ability of an index to consistently and repeatedly produce the same results with the same or different raters.[21] Validity, on the other hand, is the ability of an index to measure what it seeks to measure, i.e. treatment need, and is often measured by comparing its results to the opinions of an expert orthodontic panel, or "gold standard".[11,16,21,22] Orthodontic panels are commonly used as a "gold standard" in the orthodontic literature because their opinion of orthodontic treatment need is consistent among different geographic regions[16] and highly correlates with public opinion.[22] Validity is commonly determined using sensitivity and specificity.[16,17,23] Sensitivity describes the ability of a test to identify persons afflicted with a condition as actually *having* that condition of interest (true positive).[24] Specificity, on the other hand, is the ability of a test to identify unafflicted persons as *not having* the condition of interest (true negative).[24]

To this day, seven studies have evaluated the HLD index modifications in California, Washington, Maryland, and Illinois.[15–17,23,25–27] Only four of these compared the HLD to a gold standard.[15–17,26] Three studies validated the HLD(CalMod)[15–17] and one study validated the HLD(WA).[26] The remaining studies provided insights on index characteristics, but did not compare them to an accepted standard, and therefore do not validate the HLD indices evaluated.[23,25,27] Three studies have been conducted comparing validity and reliability of HLD(CalMod)[15–17]and one study validated the Washington HLD modification.[26] two studies evaluating the validity and reliability of the HLD(CalMod) alone,[17,25] and one study comparing the HLD(CalMod) with the Maryland HLD index, or HLD(Md).[23] The HLD(CalMod) and HLD(WA) have both been found to be highly reliable quantitative methods for identifying handicapping malocclusion characteristics and seem highly correlated with the judgements of orthodontic panels.[15–17,26] However, both indices displayed low sensitivity (25.9% - 41.7%), but high specificity (96.8% - 100.0%).[15–17,26] Simply put, these indices fail to include patients with severe malocclusion, but rule out most individuals without severe malocclusion. This seems to indicate that the cut-off scores currently in use are too stringent and limit many patients with severe malocclusions from access to orthodontic care. Studies therefore recommend lowering the cut-off score to 12 to 18.5 points, depending on the statistical model usedfor optimization of the specificity and sensitivity of the current indices.[15–17,26] In their comparative study of the HLD(CalMod) and HLD(Md) indices, Han & Davidson deduced that HLD(Md) had a higher approval rate than HLD(CalMod) (41% vs 35% respectively), and that the correlation between the two indices was relatively strong (R = 0.78).[23] In addition, they found that 70% of cases approved by the HLD(CalMod) index had an automatic qualifier and that the most common contributing factors were crowded anterior teeth for HLD(Md) and overjet for HLD(CalMod), and that HLD(CalMod) selects more severe class II malocclusions than does HLD(Md).[23] Similarly, Parker found that that 44% of cases approved by HLD(CalMod) were approved on the basis of having an automatically-qualifying exception with the remainder being approved on the basis of a score of 26 or more. It was also found that an overjet greater than 9mm comprised 26% of all approvals.[25]

Previous studies have also suggested certain modifications to the HLD index. Han & Davidson found that the HLD index fails to identify certain conditions that patients may consider to be contributing factors of a handicapping malocclusion.[23] These include localized crowding that significantly damages dental esthetics, missing teeth or spacing in the anterior dental segment, asymmetry, and dysfunctional components including posterior open bite, speech difficulties, and TMJ symptoms.[23] In addition, because malocclusion is less associated with disease states such as periodontal disease[28] than it is with negative self-image and socialization issues,[29] an improved index might include a psychosocial component to help identify those patients who would most benefit from receiving orthodontic treatment.[9,15] Interestingly, the AAO's newly proposed auto-qualifiers do not address any of the aforementioned concerns or suggestions.[19]

Finally, a careful review of the HLD indices in use today will reveal that indices leave state consultants substantial room for interpretation for orthodontic treatment need. Indeed, some indices require that eligible providers complete open-ended questions (ex: "Describe any functional concerns:") when applying for state-funded orthodontic care. This type of questionnaire may introduce persuasion bias on the part of the clinician and interpretive bias on the part of the consultant in charge of determining orthodontic treatment need.

Standardization of Indices

While the number of US states utilizing the HLD to prioritize orthodontic treatment using Medicaid funds has increased since 2006 to become the most commonly used index in the United

States as of 2015,[9] only the HLD(CalMod)—and the HLD(WA) to a lesser extent—have been validated.[15–17,26] Many of the individual state HLD indices currently in use not only vary in component measurements and cut-off scores, but they also vary in terms of eligible providers, eligible reviewers, coverage by patient age, and required records.[9] In addition, some states may or may not include an esthetic component in their evaluation of treatment need.[9] With extra-oral photographs submitted as part of the required records for some states (Washington State for example), it is possible that reviewers rely on a subjective esthetic evaluation not otherwise quantified on the index form to determine the level of treatment need. Together, these factors could have important access to care implications for patients with certain occlusal characteristics that may constitute a handicapping malocclusion in one state but not in another.

This study had two overarching aims:

1) To quantify the internal validity and reliability of all HLD indices currently in use throughout the United States by applying the indices to a sample of consecutively enrolled orthodontic patients.

2) To evaluate the implementation one the HLD indices for quality assessment by retrospectively comparing clinicians' assigned scores on the HLD(WA) pre-authorization form to the Washington state consultants' determination of treatment need using the dental records of orthodontic patients seen between April and October 2018 at five orthodontic clinics in the greater Seattle area.

AIM 1: Validity and Reliability of state-modified HLD indices

OBJECTIVES

To measure the validity and reliability of the state-modified HLD indices in the United States as of January 2019.

METHODS

Sample

Following University of Washington Institutional Review Board approval (STUDY00005449), the initial orthodontic records of 366 patients consecutively screened at the University of Washington Graduate Orthodontic Clinic between 2012 and 2013 were gathered. All personal identifiers were removed. The inclusion criteria for patient selection were 1) younger than 21 years-old, 2) presence of pre-treatment orthodontic records including physical models, extra- and intra- oral photographs, as well as panoramic and cephalometric radiographs, 3) in transitional or permanent dentition with 3 or less primary teeth remaining. The exclusion criteria were 1) age 21 or older, 2) primary or mixed dentition with more than 3 primary teeth remaining, and 3) incomplete or poor-quality pre-treatment records. 161 patients did not meet the inclusion criteria and were excluded: 71 were 21 years or older, 62 were in mixed or transitional dentition, and 28 had incomplete records or records that were out of date. Of the remaining 205 subjects, 1 subject had PAR and panel data but were missing HLD-indices data at the time of analysis. This patient was included in PAR and panel analysis, but subsequently excluded from all HLD analysis.

Peer Assessment Rating (PAR) Score Measurement

Two raters, one orthodontist (L.T.) and one second year orthodontic resident (T.C) were recruited to measure all subjects who met inclusion criteria. Both had previously received calibration training by a PAR calibrated grader. Two identical PAR rulers and as well as the pre-treatment records including physical models, extra- and intra-oral photographs, and panoramic and cephalometric radiographs were available for PAR scoring. For calibration purposes, the two raters first scored 5 randomly selected subjects from the overall sample on 2 separate occasions, each 15 days apart, until intra- and inter-correlation coefficients (ICC) of 0.80 or greater were obtained. Once the ICC reached values of 0.80 or greater, the two raters began scoring the sample of 205 patients. Scoring was performed over multiple 3-hour sessions with hourly 15 minute breaks to minimize rater fatigue. Each rater independently scored each subject's pre-treatment records and came to a consensus when scores were greater than 2 points apart. Only weighted scores were utilized. Two weeks after scoring had ended, 5 randomly-selected cases were chosen and scored a second time. The intra- (ICC(3,1)= 0.921-0.975) and inter-correlation coefficients (ICC(3,1)=0.975) were excellent.

Expert Panel

Expert Panel Recruitment

Orthodontists practicing in the Seattle-area were recruited. The inclusion criteria for taking part in the expert panel were 1) completion of an orthodontic residency program and 2) 5 or more years of orthodontic practice experience. 19 orthodontists were recruited to take part in the expert panel. 2 orthodontists were excluded from the analysis due to severely poor reliability measures.

Expert Panel Clinical Judgement Collection

Due to scheduling conflicts, it was not possible to gather all panel members at one time. Therefore, groups of 5 or less orthodontists were scheduled for scoring. An online survey platform was used to create a 100-point visual analog scale (VAS) (SurveyMonkey® SurveyMonkey.com) (Appendix 4). All panel sessions were conducted in the same standardized sequence. Panel members were equipped with 14-inch monitor laptops with a screen resolution of 1920 x 1080 pixels. A pre-evaluation script (Appendix 5) was read out loud. Panel members were provided with 10 training cases to familiarize themselves with the computer software and pace of the projection of records. Pre-treatment records were assembled on one slide per patient (Appendix 6) and projected on the same screen under the same lighting conditions at a rate of 20 seconds per slide. Pre-treatment records were presented in the order in which the patient had presented to the clinic. The order by which records were presented was the same for every panelist. Panel members were given a 10 minute break at the halfway point. 10 cases were selected by computer-randomization and repeated at the end of the presentation to measure intra-rater reliability. Inter- and intra-rater reliability was measured by calculating the intraclass correlation coefficient (ICC). Panel members were then asked to provide an indicated treatment need point at the end of the presentation using the same VAS used for scoring malocclusion severity. Finally, panel members were asked for biographical information including age, gender, residency program, date of orthodontic residency program completion, years of experience, current activity in clinical practice, and an approximate number of patients examined over the past year.

Panel members were recruited for a second rating session at least 1 month after the completion of the first session. 25 cases were randomly selected from the main sample using computer randomization software. Panel members were asked to repeat the same exercise under the same conditions, but with physical initial models in addition to pre-treatment intra- and extra-oral photographs and radiographs.

HLD Indices

Collection

The dental provider manuals as published by each state legislature were retrieved from the Medicaid websites of each of the 50 United States and the District of Columbia to identify the existence or lack of an orthodontic need index in each state. The inclusion criteria were 1) reported use of an HLD or HLD-modified index and 2) online publication of such an index. The exclusion criteria were 1) use of an index other than an HLD or HLD-modified index and 2) failure to report the use of said index on online Medicaid websites. The HLD index was selected because it is the most common single index used in the United States at the present time.[9]

A comprehensive list of automatically qualifying exceptions, component measurements, and special instructions were gathered and entered onto a single Microsoft Excel (Microsoft Inc.) spreadsheet and grouped into categories. Open-ended questions which are not ascribed a point value that counts towards the overall score on the HLD index (ex: “Describe any functional concerns,” “Describe musculature,” etc…) were omitted due to the subjective nature of these entries and because they do not contribute to the overall scores according to index instructions.

HLD Scoring

Two raters, one orthodontic faculty member (B.B.) and one orthodontic resident (T.C.) were recruited for scoring. Both raters are familiar with the HLD index and use one modification (HLD-WA) in daily orthodontic practice. The pre-treatment records included pre-treatment physical models, extra- and intra-photographs, as well as panoramic and cephalometric radiographs. A subset of 16 pre-treatment records were selected using computer randomization from the overall sample for rater calibration. Both raters independently scored all 16 sets of records and subsequently reviewed and discussed any differences until a consensus was reached. Following the training session, the raters scored all pre-treatment records in no particular order. Sessions were limited to 3 hours with hourly 10 minute breaks. Both raters examined each set of pre-treatment records together and scored each component measure in the same order for every set of records. Two identical dental rulers were used where measurements were necessary. A consensus was reached where disagreements arose.

Analysis

Intra- and inter-rater reliabilities of PAR and expert panel ratings were computed using the intra- and inter-rater correlation coefficients (ICC). Pearson correlation coefficient was used to compare PAR and panel scores. A Pearson correlation coefficient was also computed to compare and panel scores obtained with and without physical models.

To account for the lack randomization of pre-treatment records during panel scoring, autocorrelation testing was performed. A moderate autocorrelation was present, indicating that up to five preceding observations influenced panelists' score of a current observation. Autocorrelation was removed and an autocorrelation plot for the model residuals was generated. Including the lagged outcomes removed autocorrelation for the five preceding observations. Based on this model, order-adjusted panel scores were used for all HLD to panel score analyses.

To reduce each patient's 17 expert panel ratings to a single rating, order-adjusted panel scores were converted to a simple mean, sum-of-ranks, z-score standardization, proportional difference, and sum-of-votes aggregates. Pearson's r, Spearman's rho, and Kendall's tau rank correlation were used to correlate expert panel ratings to PAR scores. By all measures, all 5 aggregates were similarly and moderately positively correlated with PAR scores (Table 1).

Table 1: Correlation coefficients of expert panel aggregate measures with PAR scores

Expert Panel Aggregates	*Pearson's r*	*Spearman's rho*	*Kendall's tau*
Simple mean	0.679	0.690	0.499
Sum-of-Rank	0.694	0.697	0.505
Standardize	0.686	0.698	0.504
Proportional difference	0.689	0.702	0.506
Vote total	0.666	0.666	0.494

Total vote performed the worst (r=0.666) and sum-of-ranks performed the best (r=0.694). Sum-of-rank ratings were therefore used as the expert panel rating aggregate of choice.

Sum-of-ranks were dichotomized to "handicapping malocclusion" and "no handicapping malocclusion" to obtain the prevalence of handicapping malocclusion in the sample population.

The proportion of subjects designated as needing treatment by a majority panel vote (>8 out of 17 expert panelists) according to each panelist's own criterion was determined to be 52.9%. Based on these data, the top 52.9% of dichotomized sum-of-rank decisions were designated as having a "handicapping malocclusion." The remainder of subjects were designated as "no handicapping malocclusion."

Descriptive statistics were performed to compare state-modified HLD indices and the proposed AAO qualifying exceptions. Predicted influence of point-contributing conditions on overall HLD score was evaluated. Predicted approval and denial rates were also calculated for each state-modified index. To test the hypothesis that indices with similar predicted approval rates identified the same patients as having a handicapping malocclusion, percent agreement with other state-modified HLD indices was calculated. Percent agreement with expert panel was also computed to test the hypothesis that states which displayed the greatest agreement with other states show greater agreement with the expert panel decision for handicapping malocclusion.

Receiver Operating Characteristic (ROC) curves were generated by comparing each state-modified HLD index to the order-adjusted expert panel aggregate ratings. Areas under the ROC curve (AUROC), specificity, and sensitivity were calculated for each state at the current point threshold. Youden's J index was also calculated to generate an optimal point threshold for each state HLD.

RESULTS

Expert Panel Analysis

Expert Panel Demographics

19 orthodontists participated in the clinical judgment evaluation of patients' malocclusion severity. Two orthodontists were excluded due to poor intra-rater reliability ratings. The remaining 17 orthodontists consisted of 11 males and 6 females. Expert panel characteristics are displayed in Table 2. The mean panelist age was 54.1±12.1 years old (range: 33-78). Panelists graduated from orthodontic residency between 1970 and 2014 and 12 of 17 (70.5%) orthodontists completed orthodontic residency at the same institution (University of Washington). 15 (88.2%) panelists were involved in orthodontic practice at the time of study.

All panelists practiced in the greater Seattle area for an average of 24.1±11.7 years (range: 4.5-47 years) and reported being in current clinical practice for an average of 21.7±13.7 hours per week.

Table 2: Expert Panel Characteristics

Characteristics	*n*
Sex	
Males (n)	11
Female (n)	6
Mean Age	54.1±12.1
Mean years since orthodontic residency graduation	25.1±12.3
1-10 years (n)	2
11-20 years (n)	4`
21-30 years (n)	5
31-40 years (n)	4
40+ years (n)	2
Orthodontic programs attended (n)	5
Panelists active in orthodontic practice (%)	88.2

Expert Panel Reliability

All reliability inferences were made using raw panel scores (Table 3). The panelist inter-rater reliability was moderate (ICC(3,1) = 0.650, 95% CI, 0.603-697). The intra-rater reliability varied from moderate to excellent, ICC(3,1) = 0.686-0.974. Four raters had moderate reliability (0.50<ICC≤0.75), nine had good reliability (0.75<ICC≤0.90), and four had excellent intra-rater reliability (ICC>0.90).[30]

Table 3: Expert panel intra-rater correlation coefficients (ICC)

*Intra-rater reliability**	*ICC*	*95% CI*
Rater 1	0.974	0.900 - 0.993
Rater 2	0.956	0.832 - 0.989
Rater 3	0.931	0.744 - 0.982
Rater 4	0.907	0.676 - 0.976
Rater 5	0.884	0.533 - 0.971
Rater 6	0.881	0.610 - 0.969
Rater 7	0.872	0.586 - 0.966
Rater 8	0.871	0.583 - 0.966
Rater 9	0.850	0.523 - 0.960
Rater 10	0.847	0.290 - 0.964
Rater 11	0.833	0.357 - 0.958
Rater 12	0.775	0.357 - 0.938
Rater 13	0.769	0.344 - 0.936
Rater 14	0.737	0.217 - 0.928
Rater 15	0.735	0.268 - 0.926
Rater 16	0.715	0.181 - 0.922
Rater 17	0.686	0.182 - 0.910

ICC = intra-rater coefficient; CI = confidence interval
*ICC(3,1)—two-way mixed model, absolute, singe-measures

Influence of presence or absence of physical models on expert panel ratings

15 of the original 17 panelists rated a subset of 20 patients using the same records (extra-/intra-oral photographs and panoramic/cephalometric radiographs) and pre-treatment physical models. The mean inter-rater reliability was moderate (ICC(3,1)=0.670, 95% CI, 0.527-0.817). Expert panel ratings with and without physical models had a significant moderate positive correlation with PAR scores. The correlation was stronger for panel ratings with physical models (r(16)=.76, p<.01) than panel ratings without physical models (r(16)=.64, p<.01). Mean ratings were greater for 10 of 15 raters (66.7%) when physical models were present compared to when they were absent (mean difference=4.1).

HLD Indices Analysis

A review of HLD utilization in the United States revealed that 21 states (41.2%; n=51 including the District of Columbia) used some modification of the HLD index as of July 2019 (AK, AR, CA, DC, DE, FL, IL, MA, MD, ME, MO, MT, NJ, NM, NY, OK, RI, SC, TX, WA, WY).

Qualifying Exceptions

A summary of the qualifying exceptions is presented in Tables 4 & 5. Ten qualifying exceptions were used among all HLD-using states. No state incorporated all qualifying exceptions. Among HLD-using states, qualifying exception utilization ranged from 1 to 7 per state. MA, MO, and NY used the most, but different qualifying exceptions.

Some criteria (cleft palate deformity, craniofacial anomaly, and anterior impaction) were not listed on the HLD index of some states, despite being listed as qualifying exceptions in those states' published dental provider manuals (Appendix 1a). Cleft palate deformity was the most common qualifying exception and was used by all but one (TX) HLD-using state. Deep impinging overbite and craniofacial anomaly were the second most used qualifying exceptions, each present in 15 of 21 state HLD indices. Negative overjet relative to a skeletal Class III and severe anterior maxillary crowding (>8mm) were the least commonly used qualifying exception, each present in only one state HLD index. Some conditions (deep impinging overbite, severe traumatic deviation, crossbite of anterior teeth, and anterior impactions) were qualifying exceptions in some states and point-contributing conditions in others (Appendix 1a). MD included two point-contributing conditions (cleft palate and severe traumatic deviation) which were assigned a value equal to the point-threshold needed for case approval. The number of qualifying exceptions was moderately negatively correlated with point threshold, $r(19)=-.52$, $p<.05$.

The proposed AAO index incorporated 7 of the 10 qualifying exceptions used by HLD-using states. It also incorporated 4 additional qualifying exceptions (crowding or spacing >10mm, congenitally missing teeth, posterior impaction, lateral or anterior open bite, and anterior and/or posterior crossbite of 3+ teeth per arch).

Table 4: Qualifying exception utilization for HLD-using states (n=21)

Qualifying Exception	*# of states*	*AAO*
Cleft Palate deformity	20	✓
Deep impinging overbite	15	✓
Craniofacial anomaly	15	✓
Crossbite of individual anterior teeth	14	✓
Overjet ≥9mm or reverse overjet ≥3.5mm	12	✓
Severe traumatic deviation	12	
Anterior impaction (# of teeth)	7	✓
Negative overjet relative to a skeletal Class III	1	
Severe anterior maxillary crowding (greater than 8mm)	1	
Posterior Impaction	0	✓
Crowding or spacing of 10mm or more	0	✓
Congenitally missing teeth excluding 3rd molars	0	✓
Lateral or anterior open bite	0	✓
Anterior and/or posterior crossbite of 3+ teeth per arch	0	✓

Point-contributing conditions

The number of point-contributing conditions was not significantly correlated with predicted approval rate, $r(19)=-0.12$, $p=0.62$. Point-contributing conditions varied from state to state in the number of point-contributing conditions per index (Table 5), type of point-contributing conditions used (Appendix 1b), weighing factor (Table 6), as well as scoring instructions (Appendix 1b).

Number of point-contributing conditions per state (Table 5): There was a range of 8 to 13 point-contributing conditions per index. AR had the least (8) and MT and OK had the most (13). Most states had 9 point-contributing conditions.

Table 5: Summary of state-modified HLD indices (n=21)

State HLD	*# of QE**	*# of Point-Contributing conditions***	*Current Point Threshold*
AK	5	9	26
AR	2	8	28
CA	6	9	26
DC	6	9	15
DE	5	10	26
FL	6	9	26
IL	4	10[s]	28
MA	7	11	22
MD	2[ɣ]	12	15
ME	6	9	26
MO	7	10	28
MT	4	13	30
NJ	6	10	26
NM	4	10	30
NY	7	9	26
OK	2	13	30
RI	5	9	26
SC	2	12	35
TX	1	9	26
WA	5	9	25
WY	4	9	30
AAO	9	0	0

QE = qualifying exceptions
*Cleft palate & craniofacial anomaly counted separately; Ant and post impactions counted separately; Overjet >9mm and reverse overjet >3.5mm counted together.
**Conditions not counted as listed in HLD, but rather counted per eligible conditions counted
[s] Mx and Md crowding counted as one
[ɣ] MD's qualifying exceptions were point-contributing criteria which reached the point-threshold for automatic approval.
A comprehensive list of state HLD modifications is provided in Appendix 1a-b

Type of point-contributing condition per state (Table 6): Point-contributing conditions varied among state-modified indices. The most used were anterior open bite, overjet, overbite, and mandibular protrusion—all featured in 100% (21) of HLD-using states. Anterior and posterior ectopic eruption, anterior crowding, posterior unilateral crossbites, and labiolingual spread were used in between 16-20 HLD indices. 5 conditions were adopted by 2 to 8 states. These were posterior bilateral crossbite (DE, IL, MD, MO, MT, OK, SC, WA), severe traumatic deviation (MD, MT, NM, RI, TX, WY), anterior (AR, OK, SC) and posterior (MA, MT, SC) impaction, as well as crossbite of anterior teeth (MT, SC). Conditions including congenitally posterior missing teeth (MA), moderate and severe crowding (SC), anterior spacing (OK), midline discrepancy (OK), functional shifts (OK), habits (SC), and psychological factors (NJ) were present in only one state-modified HLD index.

Weighing factor of point-contributing conditions (Table 6): 10 conditions were weighted differently among state-modified HLD indices. 5 conditions (cleft palate, impinging overbite, severe traumatic deviation, crossbite of anterior teeth, and anterior impaction) were qualifying exceptions in some indices and point-contributing conditions in others. 8 point-contributing conditions (overjet, overbite, posterior unilateral and bilateral crossbite, severe traumatic deviation, anterior and posterior impaction, and crossbite of anterior teeth) varied in weighing factor among state-modified HLD indices. For example, overjet was multiplied by a factor of 1 in 15 states, 2 in 1 state, and 3 in one state.

Table 6: Point-contributing condition for HLD-using states (n=21)

Point-Contributing Condition	*# of states per weighing factor*									*# of states*
	"X"	*x1*	*X2*	*x3*	*x4*	*x5*	*x8*	*15pts*	*# of teeth*	*n = 21*
Overjet in mm	-	15	1	1	-	-	-	-	-	17
Overjet in mm minus 2	-	4	-	-	-	-	-	-	-	4
Overbite in mm	-	16	1	-	-	-	-	-	-	17
Overbite in mm minus 3	-	4	-	-	-	-	-	-	-	4
Ant Open bite in mm	-	-	-	-	21	-	-	-	-	21
Md protrusion in mm	-	-	-	-	-	21	-	-	-	21
Ant ectopic eruption (# of teeth)	-	-	-	20	-	-	-	-	-	20
Ant crowding (# of arches)	-	-	-	-	-	20	-	-	-	20
Post unilateral crossbite	-	-	-	-	16	2	-	-	1	19
Post ectopic eruption excluding 3rd molars	-	-	-	18	-	-	-	-	-	18
Labiolingual spread	-	16	-	-	-	-	-	-	-	16
Post bilateral crossbite	-	-	-	-	4	1	2	-	1	8
Severe traumatic deviation	-	-	-	-	1	-	-	6	-	6
Post impaction excluding 3rd molars	-	-	-	1	-	2	-	-	-	3
Ant impaction	-	-	-	2	-	1	-	-	-	3
Crossbite of ant teeth	-	-	-	-	1	-	-	-	1	2
Congenitally missing post teeth excluding 3rd molars	-	-	-	1	-	-	-	-	-	1
Cleft Palate deformity	-	-	-	-	-	-	-	1	-	1
Deep impinging overbite	-	-	-	1	-	-	-	-	-	1
Moderate crowding <6mm (# of arches)	-	-	1	-	-	-	-	-	-	1
Severe crowding >6mm (# of arches)	-	-	-	-	1	-	-	-	-	1
Ant spacing in mm	-	1	-	-	-	-	-	-	-	1
Midlines in mm	-	1	-	-	-	-	-	-	-	1
Functional shift of mandible	-	-	-	-	1	-	-	-	-	1
Psychological factors affecting development	1	-	-	-	-	-	-	-	-	1
Habits affecting arch development		-	1	-	-	-	-	-	-	1

A comprehensive list of state HLD modifications is provided in Appendix 1a-b

Point-contributing condition scoring instructions (Appendix 1a-b): Scoring instructions for qualifying exceptions and point-contributing conditions varied from state to state. For example, while mandibular protrusion (reverse overjet) is measured from the labial of the lower incisor to the labial of the upper incisor in 19 states, 2 states (DC & MA) score mandibular protrusion from the buccal groove of the first mandibular molar to the mesio-buccal cusp of the first mandibular molar. Similarly, while overjet is measured from any incisor in 13 states, 8 states score overjet only from the central incisors.

Point Threshold

Point threshold was significantly negatively correlated with predicted approval rate, $r(19)=-.46$, $p<.05$. Point-thresholds ranged from 15 (MD & DC) to 35 (SC). The mean point-threshold was 26.2±4.5 and the mode was 26, used by 9 states (Table 5).

Expert panel treatment need decisions

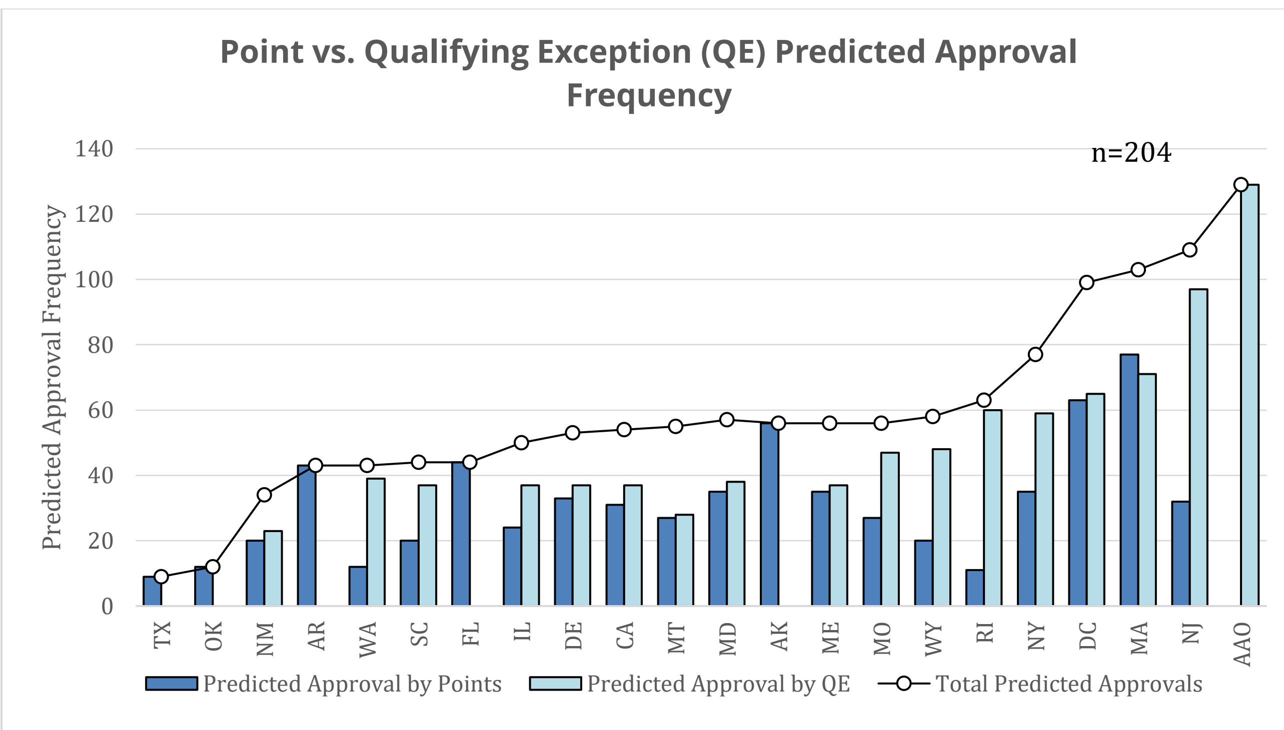
204 patient records were reviewed by an expert panel of 17 orthodontists. The mean consensus PAR score within the sample was 28.5±11.6. Dichotomized order-adjusted sum-of-rank expert panel scores were moderately positively correlated with mean consensus PAR scores, $r(202)=0.70$, $p<.001$.

Based on majority panel vote and dichotomized order-adjusted panel scores, 108 (52.94%) patients were qualified as having a handicapping malocclusion and 96 patients were deemed to have no handicapping malocclusions. Patients with a handicapping malocclusion had a statistically significant greater mean consensus PAR score (35.3±9.7) than patients with no handicapping malocclusions (21.2±8.6), $t(202)=10.93$, $p<.001$.

HLD Index Predicted Approval & Denial Rates by State

Predicted approval & denial rates: Predicted state approval and denial rates are provided in Table 7 and Figure 1. The mean predicted approval rate among HLD-using states was 29.0±14.3% (range: 4.4-53.4%) and the mean predicted denial rate was 70.9±14.2% (range: 46.3-95.6%). TX and OK had the lowest predicted approval rate at 4.4% and 5.9% respectively, while NJ and MA had the highest predicted approval rate at 53.4% and 50.5%, respectively. TX, OK, MA, and NJ were all beyond one standard deviation from the mean. The AAO index had a greater predicted approval rate (63.2%) than all state-modified HLD indices.

Figure 1: Predicted approval rates among state-modified HLD indices



Predicted utilization of qualifying exceptions: Five states did not approve any patients based on any qualifying exceptions (MD, SC, AR, OK, TX). The only qualifying exceptions for those states consisted of a combination of craniofacial anomaly, cleft palate, and severe traumatic deviation, none of which were present in the study population (Appendix 1a). Among the 16 remaining states that employed both point-contributing conditions and qualifying exceptions, 15 were predicted to approve more patients based on qualifying exceptions (mean=14.2±14.3%) than on point-contributing conditions (mean=9.3±7.2%) (Table 7). Crossbite of individual anterior teeth was the most utilized qualifying exception in 12 (57.1%) of the 21 indices (AK, CA, DC, DE, FL, IL, ME, MO, NJ, NM, OK, WA). When anterior impactions were considered a qualifying exception (NJ, MA, NY, RI, MT), they were the most utilized qualifying exception in 4 out of 5 indices, surpassing crossbite of individual anterior teeth. Anterior and/or posterior crossbite was the most utilized qualifying exception in the AAO proposed index, present in 26.0% (53) of subjects. The number of qualifying exceptions (Table 8) and the total number of approved cases per index were strongly positively correlated, $r(19)=0.72$, $p<.001$.

Table 7: Approval rates of state-modified HLD indices (n=204)

	Total Approved		*Total Denied*		*Approved on points only*		*Approved on QE only*		*Approved on points & QE*	
	n	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
AAO	129	63.2	75	36.9	0	0	129	63.1	0	0
NJ	109	53.4	95	46.3	12	5.9	77	37.9	20	9.9
MA	103	50.5	101	49.8	32	15.8	26	12.8	45	21.7
DC	99	48.5	105	51.2	34	16.7	36	17.7	29	14.2
NY	77	37.7	127	62.1	18	8.9	42	20.7	17	8.4
RI	63	30.9	141	69.0	3	1.5	52	25.6	8	3.9
WY	58	28.4	146	71.4	10	4.9	38	18.7	10	4.9
AK	56	27.5	148	72.4	19	9.4	22	10.3	16	7.9
ME	56	27.5	148	72.4	19	9.4	21	10.3	16	7.9
MD	56	27.5	148	72.4	56	27.6	0	0	0	0
MO	56	27.5	148	72.4	9	4.4	29	14.3	18	8.9
MT	55	27.0	149	72.9	27	13.3	28	13.8	0	0
CA	54	26.5	150	73.4	17	8.4	23	11.3	14	6.9
DE	53	26.0	151	73.9	16	7.9	20	9.9	17	8.4
IL	50	24.5	154	75.4	13	6.4	26	12.8	11	5.4
FL	44	21.6	160	78.3	7	3.4	24	11.8	13	6.4
SC	44	21.6	160	78.3	44	21.7	0	0	0	0
AR	43	21.1	161	78.8	43	21.2	0	0	0	0
WA	43	21.1	161	78.8	4	2.0	31	15.3	8	3.9
NM	34	16.7	170	83.3	11	5.4	14	6.9	9	4.4
OK	12	5.9	192	94.1	12	5.9	0	0	0	0
TX	9	4.4	195	95.6	9	4.4	0	0	0	0
Mean (w/o AAO)	59.2	29.0	144.8	70.9	18.9	9.3	29.0	14.2	11.4	5.6
SD	29.1	14.3	29.1	14.2	14.6	7.2	29.1	14.3	11.2	5.4
Median	55.5	27.2	148.5	72.7	14.5	7.2	25.0	12.3	10.5	5.2

Table 8: Frequency of qualifying exceptions among state-modified HLD indices (n=204)

	AK	AR	CA	DC	DE	FL	IL	ME	MD	MA	MO	MT	NJ	NM	NY	OK	RI	SC	TX	WA	WY	AAO
Cleft Palate	0	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0
Craniofacial Anomaly	-	0	0	0	-	0	-	0	-	0	0	0	0	0	0	0	-	0	0	0	-	0
Deep Impinging overbite	4	-	4	4	4	4	15	4	-	4	4	4	-	4	4	30	4	-	-	4	4	30
Crossbite of anterior teeth	22	-	22	50	22	22	22	22	-	-	22	-	75	19	22	57	22	-	-	22	22	-
Severe traumatic deviation	0	-	0	0	0	0	0	0	0	0	0	-	0	0	0	-	-	-	-	-	0	-
Overjet >9mm or reverse overjet >3.5mm	16	-	14	17	14	14	-	16	-	17	21	-	14	-	16	16	15	-	-	16	-	27
Ant. Impaction	-	-	-	-	*	-	-	-	-	25	1	25	25	*	25	-	25	-	-	-	25	36
Post. Impaction	-	-	-	-	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Negative overjet due to skeletal Cl. III	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	8	-	-
Severe Ant. Mx crowding >8mm	-	-	-	-	-	-	-	-	-	33	-	-	-	-	-	-	-	-	-	-	-	-
Crowding or spacing of 10mm or more	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	22
Ant or Post crossbite of 3+ teeth/arch	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	53
Lateral or ant open bite	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	8
Congenitally missing posterior teeth	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	9

* These qualifying exceptions were considered in provider manual but not included in HLD index; Impactions were therefore not counted.

Predicted percent agreement among states: We computed the mean percent agreement among state-modified HLD indices to test whether states were approving the same or different patients (Figure 2; Table 9). States displayed a mean percent agreement of 79.1±5.3%. DE had the highest mean predicted percent agreement (85.2%) while NJ had the lower lowest mean predicted percent agreement (66.7%) with other states. Only one pair of states were predicted to have perfect agreement (ME & AK). TX and NJ had the poorest predicted state agreement (50.5%). States which were predicted to approve the same number of patients did not display perfect between-state percent agreement (range: 78.4-100.0%).

Figure 2: Mean predicted percent agreement among state-modified HLD indices

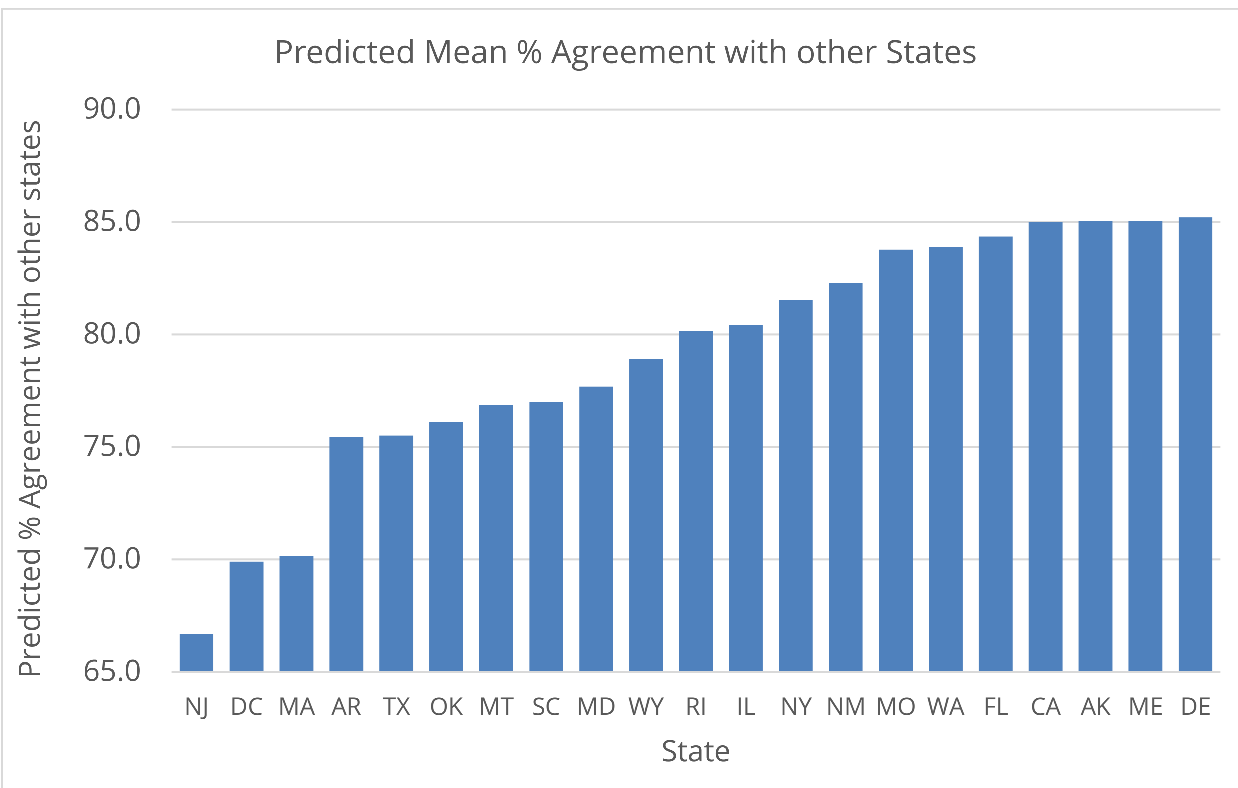


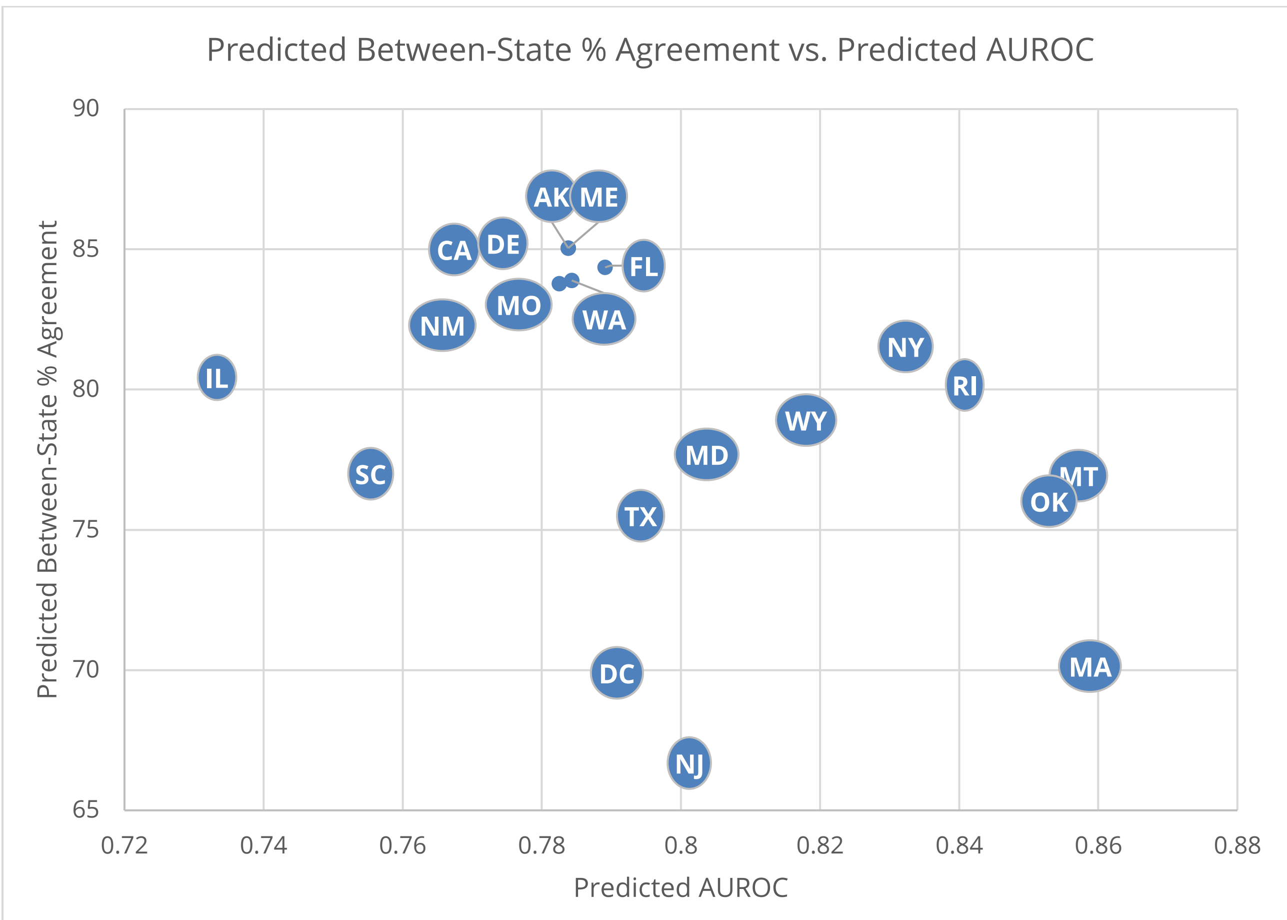
Table 9: Percent agreement among state-modified HLD indices (n=204)

	AK	AR	CA	DE	FL	IL	ME	MD	MA	MO	MT	NJ	NM	NY	OK	RI	SC	TX	WA	WY	DC	Total
AK	NA	80.9	99.0	98.5	94.1	88.7	100.0	82.8	72.5	95.1	76.5	71.1	88.7	89.7	77.0	81.9	81.4	76.5	92.6	79.9	73.9	-
AR	80.9	NA	79.9	80.4	77.9	76.5	80.9	81.4	63.2	79.9	70.1	55.9	75.5	71.6	82.4	70.6	90.7	81.9	78.4	66.7	64.5	-
CA	99.0	79.9	NA	99.5	94.1	89.7	99.0	81.9	72.5	95.1	75.5	70.1	89.7	88.7	77.9	81.9	81.4	77.5	93.6	78.9	73.9	-
DE	98.5	80.4	99.5	NA	94.6	90.2	98.5	82.4	73.0	95.6	75.0	69.6	90.2	88.2	78.4	82.4	81.9	77.9	94.1	79.4	74.4	-
FL	94.1	77.9	94.1	94.6	NA	89.7	94.1	79.9	67.6	94.1	78.4	65.2	93.6	83.8	81.9	86.8	78.4	82.4	97.5	82.8	70.0	-
IL	88.7	76.5	89.7	90.2	89.7	NA	88.7	77.5	64.2	88.7	74.0	59.8	90.2	78.4	80.4	79.4	76.0	79.9	89.2	81.4	66.0	-
ME	100.0	80.9	99.0	98.5	94.1	88.7	NA	82.8	72.5	95.1	76.5	71.1	88.7	89.7	77.0	81.9	81.4	76.5	92.6	79.9	73.9	-
MD	82.8	81.4	81.9	82.4	79.9	77.5	82.8	NA	69.1	80.9	75.0	64.7	79.4	77.5	78.4	72.5	83.8	77.0	79.4	74.5	72.9	-
MA	72.5	63.2	72.5	73.0	67.6	64.2	72.5	69.1	NA	71.6	74.5	77.0	63.2	82.8	55.4	77.9	69.6	53.9	68.1	74.0	79.8	-
MO	95.1	79.9	95.1	95.6	94.1	88.7	95.1	80.9	71.6	NA	75.5	68.1	88.7	85.8	77.0	82.8	80.4	76.5	92.6	79.9	71.9	-
MT	76.5	70.1	75.5	75.0	78.4	74.0	76.5	75.0	74.5	75.5	NA	69.1	81.9	86.8	77.9	86.8	73.5	77.5	76.0	90.7	66.5	-
NJ	71.1	55.9	70.1	69.6	65.2	59.8	71.1	64.7	77.0	68.1	69.1	NA	59.8	81.4	52.0	73.5	59.3	50.5	64.7	71.6	79.3	-
NM	88.7	75.5	89.7	90.2	93.6	90.2	88.7	79.4	63.2	88.7	81.9	59.8	NA	78.4	87.3	81.4	77.0	87.7	92.2	85.3	67.0	-
NY	89.7	71.6	88.7	88.2	83.8	78.4	89.7	77.5	82.8	85.8	86.8	81.4	78.4	NA	67.6	92.2	74.0	66.2	82.4	90.2	75.4	-
OK	77.0	82.4	77.9	78.4	81.9	80.4	77.0	78.4	55.4	77.0	77.9	52.0	87.3	67.6	NA	73.5	81.9	98.5	83.3	77.5	57.1	-
RI	81.9	70.6	81.9	82.4	86.8	79.4	81.9	72.5	77.9	82.8	86.8	73.5	81.4	92.2	73.5	NA	74.0	73.0	87.3	92.2	71.4	-
SC	81.4	90.7	81.4	81.9	78.4	76.0	81.4	83.8	69.6	80.4	73.5	59.3	77.0	74.0	81.9	74.0	NA	82.4	78.9	68.1	66.0	-
TX	76.5	81.9	77.5	77.9	82.4	79.9	76.5	77.0	53.9	76.5	77.5	50.5	87.7	66.2	98.5	73.0	82.4	NA	82.8	76.0	55.7	-
WA	92.6	78.4	93.6	94.1	97.5	89.2	92.6	79.4	68.1	92.6	76.0	64.7	92.2	82.4	83.3	87.3	78.9	82.8	NA	81.4	70.4	-
WY	79.9	66.7	78.9	79.4	82.8	81.4	79.9	74.5	74.0	79.9	90.7	71.6	85.3	90.2	77.5	92.2	68.1	76.0	81.4	NA	68.0	-
DC	73.9	64.5	73.9	74.4	70.0	66.0	73.9	72.9	79.8	71.9	66.5	79.3	67.0	75.4	57.1	71.4	66.0	55.7	70.4	68.0	NA	-
Mean	85.0	75.5	85.0	85.2	84.4	80.4	85.0	77.7	70.1	83.8	76.9	66.7	82.3	81.5	76.1	80.2	77.0	75.5	83.9	78.9	69.9	**79.1**
SD	9.1	8.0	9.3	9.2	9.4	9.0	9.1	4.9	7.2	8.7	5.8	8.3	9.5	7.5	10.7	6.5	7.0	11.1	9.1	7.1	6.1	**5.3**

Green= Highest % agreement, yellow= moderate % agreement; red= poorest % agreement

Predicted percent agreement with expert panel: We tested the hypothesis that states which displayed the greatest agreement with other states would show greater agreement with the expert panel decision for handicapping malocclusion (Figure 3). The hypothesis was rejected—there was no statistically significant correlation between these two variables, r(19) = -0.25, p=0.81.

Figure 3: Predicted between-state percent agreement vs. predicted HLD AUROC



Predicted influence of point-contributing conditions on overall HLD score: Percent influence of point-contributing conditions on overall HLD index scores were evaluated 1) as a mean when the condition was present and 2) as an overall mean regardless of whether the condition was present (Table 10). Mandibular protrusion, anterior open bite, and anterior ectopic eruption had the greatest influence (41.8%, 41.5%, and 41.8% respectively) on overall HLD index score, when present. When state HLD indices included point-contributing criteria that were not common among other indices, these tended to have low to moderate influence on total HLD score (0-27.8%). Of these, bilateral crossbite, impactions, and anterior spacing were most influential. Only overbite and overjet had consistent influence on total HLD index score (30.6-33.9%) regardless of whether the condition was present or not.

Table 10: Mean percent influence of point-contributing conditions on overall score
Percent influence of point-contributing conditions on total HLD index score was evaluated 1) as a mean when the condition was present and 2) as an overall mean regardless of whether the condition was present.

	Mean % overall weight	Mean % weight if condition present	Mean % Difference	# of States with condition
Overjet equal to or less than 9mm	30.6%	33.9%	3.3%	21
Overbite in mm	22.9%	26.7%	3.8%	21
Mandibular protrusion in mm	4.6%	41.8%	37.2%	21
OJ or md protrusion	31.1%	31.3%	0.2%	1
Open bite in mm	4.1%	41.5%	37.4%	21
Anterior Ectopic eruption only	1.1%	19.3%	18.2%	1
Posterior ectopic eruption	1.5%	18.5%	17.0%	17
Anterior crowding only	6.0%	41.8%	35.8%	1
Ectopic OR anterior crowding	20.7%	35.8%	15.1%	19
Labiolingual spread in mm	14.5%	23.3%	8.8%	16
Posterior unilateral crossbite	2.8%	20.4%	17.4%	18
Anterior crossbite	5.3%	14.2%	9.1%	2
Bilateral crossbite	1.0%	27.8%	26.8%	1
Posterior crossbite	3.3%	8.3%	5.0%	1
Anterior or posterior impactions	3.0%	20.5%	17.5%	1
Anterior impaction	1.7%	17.9%	16.2%	1
Posterior impactions or congenitally missing posterior teeth	2.3%	22.1%	19.8%	1
Posterior Impaction	1.4%	20.3%	18.9%	1
Severe Traumatic Deviations	0.0%	0.0%	0.0%	5
Deep impinging overbite	0.3%	15.8%	15.5%	1
Anterior spacing	4.8%	22.5%	17.7%	1
Functional shift of mandible	0.2%	13.6%	13.4%	1
midlines	7.2%	10.3%	3.1%	1
Moderate crowding	2.9%	8.4%	5.5%	1
Severe crowding	4.4%	16.6%	12.2%	1
Habits	0.1%	7.7%	7.6%	1
Cleft Palate	0.0%	0.0%	0.0%	1

HLD-decisions vs. Expert Panel Scores

State-modified HLD decisions were compared to dichotomized order-adjusted sum-of-rank expert panel decisions. The areas under the receiver operating characteristic (AUROC) curves, a measure of the diagnostic value of each index, were subsequently computed. A perfect diagnostic test would have an area under the curve (AUC) of 1.00. AUROCs for each HLD-using state are represented in Figure 4 and Table 11 at each state's current point threshold. A complete list of state-modified HLD index ROC curves can be viewed in Appendix 2.

The AUROCs were moderately high (mean=0.807±0.041). AR had the lowest AUROC (0.703), and MA had the highest AUROCs (0.875). The specificity, or ability for HLD indices to exclude patients without medical necessity, was high (mean=92.9±5.0%) and displayed a modest range (range=79.5-100%). NJ (79.5%) had the lowest specificity and OK & TX (100%) had the highest specificity. Sensitivity, or the ability for HLD indices to correctly identify patients with medical necessity, was low (mean=41.6±17.6%) and displayed a wide range (7.5-77.5%). TX (6.67%) and OK (9.17%) had the lowest sensitivity while MA (77.5%) and NJ (76.7%) had the highest sensitivity. The mean sensitivity and specificity margins of error were less than 10%. The positive predictive value for state-modified HLD indices was very high, ranging from 83.8-100%. The negative predictive value on the other hand was moderate, ranging from 42.8-73.3%. There was no correlation between a state-modified HLD index's AUROC and the total number of subjects accepted by state, r(19)=.24, p=0.304, the point threshold, r(19)=-.09, p=0.93, or the number of qualifying exceptions, r(19)=.29, p=.77.

The performance of the proposed AAO index varied considerably from that of state modified HLD indices. Compared to HLD indices, it had a high sensitivity (88.8%) and a moderate specificity (65.6%). Its positive predictive value (74.4%) was also lower than all HLD indices and its negative predictive value was greater (85.5%) than all HLD indices. 85.4% (82 of 96) of patients deemed not have a handicapping malocclusion by the expert panel were approved by 2 or less state-modified HLD indices (Figure 5). However, only 39.3% (42 of 107) of patients deemed to have a handicapping malocclusion by the expert panel were approved by more than 50% (11 or more) of state-modified HLD indices. These results are consistent with earlier findings that state-modified HLD indices generally have high specificity and low sensitivity.

Figure 4: Area under the ROC (AUROC) for HLD scores vs Rank-sum Panel Scores

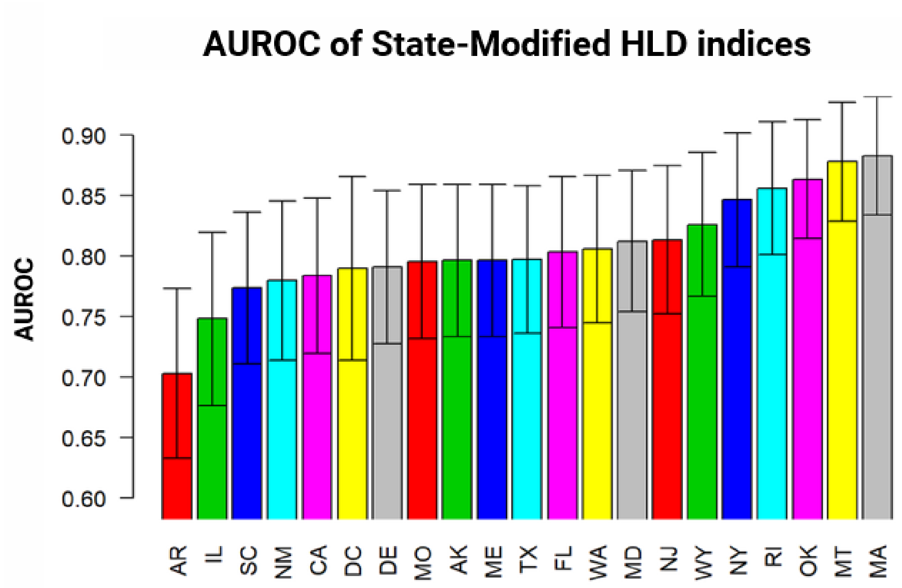


Figure 5: Comparison of state-modified HLD and expert panel decisions

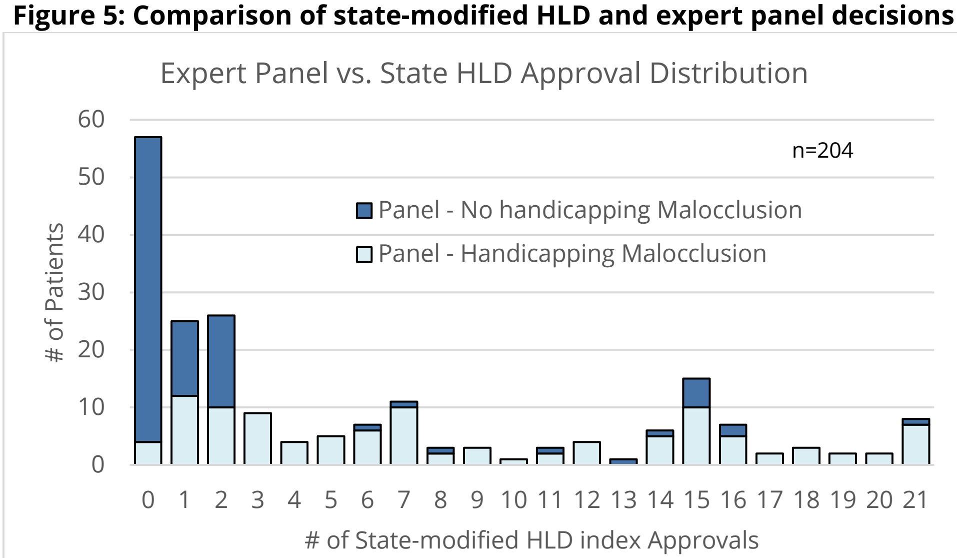


Table 11: State-modified HLD index validity at current point thresholds (n=204)

				Sensitivity		*Specificity*		*(+) Predictive value*	*(-) Predictive value*
State	*Point Threshold*	*AUROC*	*95% CI*	*%*	*Margin of Error*	*%*	*Margin of Error*	*%*	*%*
MA	22	0.883	0.834-0.932	77.5	7.8	89.2	7.3	91.2	73.3
MT	30	0.878	0.829-0.927	43.3	9.2	96.4	4.7	94.6	54.1
OK	30	0.864	0.814-0.913	10.0	5.8	100.0	2.2	100.0	43.5
RI	26	0.856	0.802-0.911	47.5	9.3	92.8	6.2	90.5	55.0
NY	26	0.847	0.791-0.902	59.2	9.1	92.8	6.2	92.2	61.1
WY	30	0.826	0.767-0.886	43.3	9.2	92.8	6.2	89.7	53.1
NJ	26	0.813	0.752-0.875	76.7	7.9	79.5	9.2	84.4	70.2
MD	15	0.812	0.754-0.871	44.2	9.2	96.4	4.7	94.6	54.4
WA	25	0.806	0.745-0.867	31.7	8.7	94.0	5.8	88.4	48.8
FL	26	0.803	0.741-0.866	32.5	8.7	94.0	5.8	88.6	49.1
TX	26	0.797	0.736-0.858	7.5	5.1	100.0	2.2	100.0	42.8
AK	26	0.796	0.734-0.859	42.5	9.2	94.0	5.8	91.1	53.1
ME	26	0.796	0.734-0.859	42.5	9.2	94.0	5.8	91.1	53.1
MO	28	0.796	0.732-0.859	40.8	9.1	91.6	6.6	87.5	51.7
DE	26	0.791	0.728-0.854	40.0	9.1	94.0	5.8	90.6	52.0
DC	15	0.790	0.714-0.866	69.2	8.6	80.7	9.0	83.8	64.4
CA	26	0.784	0.720-0.848	40.8	9.1	94.0	5.8	90.7	52.4
NM	30	0.780	0.714-0.845	24.2	8.0	94.0	5.8	85.3	46.2
SC	35	0.774	0.711-0.836	35.0	8.9	97.6	4.1	95.5	50.9
IL	28	0.748	0.677-0.820	33.3	8.8	88.0	7.6	80.0	47.7
AR	28	0.703	0.633-0.773	32.5	8.7	95.2	5.3	90.7	49.4
AAO	-	-	-	88.8	6.1	65.6	11.6	74.4	85.3
Mean	26.2	0.807	0.746-0.868	41.6	8.5	92.9	5.8	90.5	53.6
SD	4.5	0.041	0.048-0.036	17.6	1.1	5.0	1.7	4.9	7.8

QE=Qualifying exception; CI = confidence interval

Current vs. Optimal Point Threshold

Youden's J statistic was calculated for all state-modified HLD indices to provide optimal point thresholds by maximizing the difference between the true positive rate (TPR) and false positive rate (FPR). There was a statistically significant difference between the mean optimal (18.9±.49) and current (26.2±4.4) point thresholds, t(19)=4.93, p<.0001, as well as mean current and optimal sensitivities, t(19)=-6.62, p<.0001, and specificities, t(19)=6.23, p<.0001 (Table 12). The mean sensitivity for all state-modified HLD indices increased from 40.0% at the current thresholds to 84.4% at the optimal thresholds (p<.0001), while the mean specificity decreased from 93.5% to 68.8% (p<.0001) (Table 12 & 13). All calculated optimal point thresholds were lower than the current point thresholds (mean difference=7.3±3.7) (Table 13). TX (point dif=16), OK (point dif=14), FL (point dif=14), and RI (point dif=11) had the greatest point difference and were beyond one standard deviation from the mean. MA (point dif=1), AR (point dif=2), and NJ (point dif=3) had the lowest current to optimal point threshold differential.

Table 12: Current vs. optimal point threshold summary

	Current	*Optimal*	*P-value**
Mean point threshold	26.2±4.4	18.9±4.9	<.0001
Mean sensitivity (%)	40.0±16.7	68.8±10.1	<.0001
Mean specificity %)	93.5±4.2	84.4±5.0	<.0001

*Two-sample Student's t-test assuming equal variances

Table 13: State-specific current vs. optimal point threshold

		Current Point Threshold			*Optimal Point Threshold*			
State	*AUROC*	*Sens (%)*	*Spec (%)*	*Point threshold*	*Point threshold*	*Spec (%)*	*Sens (%)*	*Point threshold Difference*
MA	0.883	77.5	89.2	22	21	89.2	80.0	1
MT	0.878	43.3	96.4	30	21	85.5	78.3	9
OK	0.864	10.0	100.0	30	16	80.7	79.2	14
RI	0.856	47.5	92.8	26	15	79.5	83.3	11
NY	0.847	59.2	92.8	26	21	89.2	72.5	5
WY	0.826	43.3	92.8	30	21	83.1	72.5	9
NJ	0.813	76.7	79.5	26	23	78.3	81.7	3
MD	0.812	44.2	96.4	15	10	72.3	80.8	5
WA	0.806	31.7	94.0	25	16	88.0	64.2	9
FL	0.803	32.5	94.0	26	15	88.0	63.3	11
TX	0.797	5.8	100	26	10	83.1	71.7	16
AK	0.796	42.5	94.0	26	21	89.2	60.8	5
ME	0.796	35.0	94.0	26	21	89.2	60.8	5
MO	0.796	40.8	91.6	28	19	83.1	70.0	9
DE	0.791	40.0	94.0	26	19	86.8	64.2	7
DC	0.790	43.9	93.1	15	11	75.0	74.2	4
CA	0.784	40.8	94.0	26	19	85.5	63.3	7
NM	0.780	24.2	94.0	30	21	89.2	57.5	9
SC	0.774	35.0	97.6	35	31	90.4	51.7	4
IL	0.748	33.3	88.0	28	19	79.5	70.0	9
AR	0.703	32.5	95.2	28	26	86.8	45.0	2
Mean	0.807	40.0	93.5	26.2	18.9	84.4	68.8	7.3
SD	0.041	16.7	4.2	4.4	4.9	5.0	10.1	3.7

DISCUSSION

Parker reported in 1998 that while the HLD was gaining momentum as the treatment need index of choice, the subtle differences among each state-modified HLD was cause for concern.[14] Since he published his findings, only four HLD validation studies evaluating three state-modified HLD indices have been published in the orthodontic literature.[16,17,23,27] The remaining HLD indices lack supporting evidence for their use as treatment need indices for identifying handicapping malocclusions among

Medicaid recipients. To our knowledge, this is the first study to compare and compute the internal validity and reliability of all existing HLD indices in the United States.

The differences among state-modified indices under the categorical "HLD" designation were widespread. While many state Medicaid agencies have adopted the California HLD modification following its development and validation in the 1990s, almost all have subsequently modified it.[14,25,31] A direct comparison between HLD studies is only possible if the indices have not undergone substantial qualitative change between study publication dates and if the indices were validated using similar methods. Of the 21 indices validated in this study, only two have previously been evaluated against an expert panel: HLD(CA), HLD(MD).[15–17,23,25] There is published data on the validity of HLD(IL) and HLD(WA), but due to methodological differences, these could not be compared with the results from the present study.[26,27] Since the HLD(CA) has been most studied in the orthodontic literature,[14–17,23,25,31] it serves as a good basis for comparison against other state-modified HLD indices. Our results reveal that Medicaid state agencies have added, removed, or altered the number, weighing factor, and definitions of qualifying conditions, point-contributing conditions, and overall point threshold. Only AK, DE, FL, and ME have adopted the HLD(CA) modification without major changes to its component criteria or overall point threshold of 26. Still, close examination of these modifications reveals minor changes in the definitions of the component criteria. For example, the Florida modification subtracts 2mm and 3mm from overjet and overbite, respectively. The continued practice of modifying state HLD indices without supportive validation studies is a concern because it may result in screening instruments which are inconsistent with the Medicaid statutes that limited tax-payor funds be allocated to patients with orthodontic medical necessity.[32]

Overall, our predicted approval rates appear more conservative than what has been published in the orthodontic literature.[23,25–27] In this study, we found that the mean predicted approval rate was 29.0% and that the range was 4.4-53.4% (Table 7). HLD(CA), specifically, had an overall predicted approval rate of 26.5%. This comes close to the HLD(CA) predicted approval rate found by Han and Davidson (35%),[23] but contrasts sharply with the low approval rate found by Cooke and colleagues (7.2%)[17] and the high approval rate reported by Parker (56%).[25] Given the relative uniformity of scoring conditions among these studies, approval rate differences are likely due to sampling differences. For example, Cooke and colleagues included post-treatment models in their sampling population. This fact likely decreased the overall sample malocclusion severity and predicted approval rate. Similarly, Parker's 1999 sample may have suffered from sampling bias as sampling in that study was neither random nor consecutive.[25] Dr. Parker also had a significant conflict of interest as the main architect to the HLD(CA) and the contracted orthodontic consultant to the state of California during HLD(CA)'s development.[14]

Similarly, our study results predicted an HLD(MD) approval rate of 27.5% compared to Han and Davidson's 40.6%.[23] Unlike the HLD(CA), the HLD(MD) has undergone minor modifications since Han and Davidson published their results. Namely, a posterior crossbite component has been added. It is unclear why the addition of this component did not increase the overall approval rate in our results compared to those of Han and Davidson. No data was present in the orthodontic literature regarding all other state-modified HLD indices.

The range in approval and denial rates in our sample appears to be largely the result of outlier HLD indices. 14 out of 21 states had an approval rate ranging from 21.1-30.9%. When the five states with the highest and lowest approval rates were included, the range of approval rates increased to 4.4-53.7%. In many of these outlier states, significant modifications were made to the number of qualifying exceptions, point-contributing conditions, or point-threshold as compared to the HLD(CA) index (Appendix 1a-b). The number of qualifying exceptions present in an index seems especially important to its predicted approval rate, as evidenced by the strong positive correlation between these two variables. The low predicted approval rates observed in the HLD(OK) and HLD(TX), for example, appear to be the result of a lack of qualifying exceptions.

Similarly, drastic modification to the point threshold can also result in drastic changes to the overall approval rate. In 2000, Parker noted that an increase in HLD score results in an exponential decrease in approval rates.[31] While our results support this claim and a moderate negative correlation was found between point-threshold and overall predicted approval rates, the addition or removal of conditions can counterbalance the effects of point-threshold on approval rates. For example, while MD and DC both had the lowest point-thresholds (15), HLD(DC) had a predicted approval rate double that of HLD(MD). This appears to be in large part due to HLD(DC)'s use of qualifying exceptions, which are absent from the HLD(MD).

Overall, it appears that the complex process by which malocclusions are screened by an HLD index is highly dependent on the morphological characteristics that index seeks to identify. Specifically, indices are highly sensitive to changes in the number and definitions of qualifying exceptions first and foremost. The value of the point-threshold was the second most important factor related to predicted approval rates.

Our findings regarding state-modified HLD indices also apply to the proposed AAO qualifying exceptions index. The proposed AAO index consists of nine qualifying exceptions. Any patient who presents with at least one qualifying exception would gain approval to publicly-funded orthodontic care under this index. Our results indicate that the AAO proposed index had the greatest predicted approval rate among HLD modified indices (63.2%). In contrast, a 2018 unpublished study by Oppenhuizen found a lower predicted approval rate of 47.5%.[33] However, we must exercise caution when comparing our results with those of Oppenhuizen. The proposed AAO index used in his study appears mildly modified from that used in this study and published in a 2017 AAO bulletin.[19] Specifically, it excludes the "crowding or spacing of 10mm or more" component included in the present study, and several modifications were made to some of the included qualifying exceptions. While removing the "crowding and spacing of 10mm or more" did lower the predicted approval rate from 63.2% to 58.8% in our study, this rate still exceeded that of Oppenhuizen by over 10%. It is unclear why this discrepancy exists, but it is consistent with the trend that our results appear generally more conservative than those of previous studies.

Overall, the proposed AAO index seems to fit the trend that qualifying exceptions are strongly positively correlated with predicted approval rate. If the AAO's aim is to qualify 10% of the general population, as stated in a 2017 AAO bulletin, a qualifying exception based index does not seem to support this goal according to both Oppenhuizen and this study's results.[19,33] However, preliminary observations of state expenditure in Nevada indicate otherwise. In 2018, Nevada piloted the AAO's proposed index and later published its expenditure pre- and post- index implementation.[34] A decrease of 28% in expenditure on medically necessary orthodontic care was noted when comparing the 4 months before and after implementation.[33,34] A rigorous study of the Nevada Medicaid agency's implementation of its new index against the opinion of an expert panel would be necessary to explain these apparent discrepancies.

Despite heterogeneity among predicted approval rates, it would be expected that indices with similar approval rates would approve the same patients. However, this did not seem to always be the case (Table 9). While AK, MD, ME, MO were predicted to approve 56 patients each, their percent between-state agreement ranged from 80.9% (MD & MO) to 100% (AK & ME). Similarly, while AR & WA and FL & SC were predicted to approve the same number of patients (43 and 44 patients, respectively), both pairs of states displayed 78.4% agreement. Much of the disagreement occurs between opposite outlier states. HLD(NJ), for example, approved the greatest number of patients while HLD(TX) had the lowest approval rate among all HLD indices. Yet, both states had the lowest percent mean agreement with other states, the worst of which being with each other (% agreement = 50.5%). This is concerning because indices were developed to provide uniformity to the process of identifying handicapping malocclusions. If states do not agree on what constitutes orthodontic medical necessity using what should be the same HLD index, they are either using different indices or have not reached a consensus on what constitutes a handicapping malocclusion.

While predicted approval rate and between-state agreement provide important information on what proportion of the orthodontic patient population can be expected to gain access to orthodontic care, they do not shed light on whether the indices are correctly identifying patients with handicapping malocclusions. To this end, comparison of an orthodontic treatment need index to the opinions of an expert panel is necessary. When we applied this methodology, we found that HLD-indices had generally moderate validity or agreement with the orthodontic panel (measured as AUROC), low ability to identify patients with handicapping malocclusion (sensitivity), and high ability to identify patients without handicapping malocclusion (specificity). These findings follow a general trend that is well-documented in the orthodontic literature.[15–17] There were, however, some differences in the values obtained. For example, while an AUROC of 0.784 was obtained for HLD(CA) in the present study, previous studies reported AUROCs of 0.940 and 0.961 for this index (Table 11).[15,16] Similarly, while a 0.260 HLD(CA) sensitivity was previously reported,[15,17] our study found a sensitivity nearly twice that value (0.408). The discrepancies between this study and previous ones are likely rooted in the performance of the expert panel. The moderate inter- and intra-rater coefficients obtained are either due to rater confusion regarding the task at hand or measurement error related to the nature of the scale used. In either case, systematic errors likely had an effect on measures of validation investigated in this study.

Interestingly, no statistically significant correlation was found between AUROC and the number of patients accepted per index. In fact, state-modified HLD indices with the lowest predicted approval rates (HLD(OK) and HLD(TX) for example) had AUROC values greater than those of HLD(CA), which had a moderate predicted approval rate. These states also had 100% positive predictive values, indicating that 100% of patients accepted by these indices were also qualified as having handicapping malocclusions by the panel. Concluding that these indices are effective at identifying all patients with handicapping malocclusions would be erroneous, however. These same state-HLD indices had the lowest sensitivities (7.5-10.0%), highest specificities (100%), and lowest negative predictive values (42.8-43.5%) among all HLD indices evaluated in this study, suggesting that HLD(OK) and HLD(TX) are highly restrictive and deny a significant proportion of patients with handicapping malocclusions.

Despite its simplicity, the index with the best overall performance was the proposed AAO index. It displayed a high sensitivity, moderate specificity, and moderate positive and negative predictive values. However, this strong performance comes at a cost of approving a high proportion of the orthodontic patient population. While this might lead to improved access to care among Medicaid orthodontic patients, it may not be realistic for Medicaid agencies which operate on a limited budget.

One major challenge for all validation studies, including this one, is the lack of consistent methodology for determining a dichotomous cut point for medical necessity within a sample. Medicaid legislation is intentionally vague regarding medical necessity, leaving the decision on what constitutes medical necessity to states and their contracted consultants.[35] Data from NHANES III as well as Breistein and colleagues estimated that the rate of treatment need in the general population ranged from 11.9-22.6% depending on the study population.[20,36] While there is no published data regarding the rate of orthodontic medical necessity in an orthodontic patient population, it would be expected to be greater than that reported in the general population. Previous studies have used a variety of methods for dividing samples into elective treatment and medical necessity categories. Some authors[15,16] asked raters to provide an indicated treatment need point on an ordinal scale, while others[17] used an ordinal scale with an inherent “treatment essential” anchor point. According to these methods, orthodontic medical necessity was estimated at 17.6%[17] and 64%[16] in the sample orthodontic patient populations. While 17.6% appears low for an orthodontic population considering the rates observed in the general population, 64% seems more plausible. Through a different protocol, our study established that 52.94% of the sample to have a handicapping malocclusion. While this estimate appears within reason, it is likely that the closer a patient is to the cut point, the more uncertainty there is about their malocclusion severity categorization. This could introduce error in the calculations of state-modified HLD screening performance. However, it is also a reflection of what occurs in daily practice when using an index which includes both a continuous measure of malocclusion severity and a dichotomous point threshold, like the HLD.

In orthodontics, researchers and public health officials should seek to develop indices which maximize the correct identification of handicapping malocclusions, even if it means accepting patients who do not have orthodontic medical necessity. This entails creating and implementing indices with high sensitivities and moderate specificities. The proposed AAO index seems to fit these criteria better than any existing state-modified HLD index at the moment. However, given that state Medicaid agencies are not likely to restructure the process by which they screen orthodontic patients, an approach which simultaneously maintains and improves upon existing HLD indices is needed. A new index should be developed with certain criteria in mind: It should be fair, reliable, highly sensitive, and reasonable for states to implement, adjust, and afford. The AAO proposed index may not meet these criteria.

As previous investigators noted, the simplest way of improving existing HLD indices is by altering their overall point thresholds.[15–17,23] The methodology by which we sought to compute an optimal point threshold for each HLD index was based on the premise that sensitivity and specificity should be maximized for proper identification of handicapping malocclusions. When we applied this procedure to state-modified HLD indices, all resulting optimal point threshold points were lower than the current ones. Our finding that lowering the overall point threshold results in greater index validity in all cases agrees with those of previous studies.[16,17] In fact, Cooke et al recommended that HLD(CA) lower its point threshold from 26 to 18.5 points for a targeted sensitivity and specificity of 55.6% and 92.9%, respectively.[17] Our calculations yielded a similar optimal point threshold of 19 points for resulting sensitivities and specificities of 63.3% and 85.5%, respectively (Table 13).

The wide range of optimal to current point threshold difference speaks to the need for optimization of state-modified HLD indices. It is interesting to note that HLD indices set to the calculated optimal point threshold would yield the same validity, as measured by sensitivity and specificity, as that of the proposed AAO index in its current form. But given the findings presented in this study, a case can be made for maintaining the current HLD indices.

Firstly, this study provides scientific evidence of the strengths and weaknesses of each state-modified HLD index. If aware of them, state consultants could use their levying powers to override the index when it falls short of identifying what the orthodontic opinion would consider a handicapping malocclusion. In addition, there should be less incentive to change the components of an index if it has been thoroughly validated. Changing it would only result in greater ambiguity as to its screening ability.

Secondly, the AAO index lacks any published validation study involving an expert panel. If adopted prior to validation, the proposed AAO index would perpetuate the cycle of unvalidated indices being used and modified without supporting scientific evidence.

Third, the HLD index allows Medicaid agencies and public officials the liberty to alter state expenditure on medically necessary orthodontic care in any given year by simply adjusting the point threshold, rather than modifying its component measures. This is not true of the proposed AAO index which cannot be modified under its present form without adding, removing, or altering one of its nine qualifying exceptions. This is perhaps its biggest drawback to the proposed AAO index from the perspective of a public administration agency.

Limitations and Future Studies

This study has several limitations. A significant limitation was that the orthodontic panel inter-reliability was only moderate (ICC(3,1)=0.650) and there was a wide range of intra-rater reliability (ICC(3,1)=0.686-0.974). Presence of physical models had a negligible improvement effect on overall inter-rater reliability (ICC(3,1)=0.670. It is likely that the variation in inter and intra-rater reliability contributed to the observed discrepancy between AUROC, sensitivity, and specificity measures of state-modified HLD indices compared to findings from other studies. Interestingly, similar validation studies report much greater inter- and intra-reliability.[15,17,23] A possible reason for this discrepancy is the nature of the scale used for determining malocclusion severity by the expert panel. The scale used in the present study was a continuous 100-point visual analog scale with no tick marks and two anchors—ideal to most severe (Appendix 4). This type of scale is regularly used in pain,[37] dental esthetics,[38–41] and malocclusion severity[42] research. It was specifically selected in this study to provide uniformity among the indices of comparison (PAR and HLD scoring is done on a continuous scale in both instances). Similar orthodontic index validation studies did not use a visual analog scale (VAS) and instead elected for ordinal scales.[15–17] While one study did not report the panel expert panel reliability,[17] the other two obtained an interrater reliability of 0.81-0.835 and an intra-rater kappa score of 0.896-0.91, all of which are considerably greater than the values computed in the present study. This could be due to several reasons. VAS scales have been reported to incur 20% measurement error in some studies,[37] an observation which is likely related to the lack of anchor references between the two terminal anchors. It is also possible that individual raters tended to gravitate toward different parts of the scale depending on individual assessment biases and the raters' personal views on publicly-funded orthodontic care. To overcome this issue, we conducted several tests, include a transformation to Z-scores and tested the correlation of panel scores to PAR scores. Z-score transformation did not improve the correlation or reliability measures, indicating that this may not have been the leading factor in observed ICC values. Another possibility is that ordinal scales limit the range of possible values and therefore make it more plausible for inter- and intra-reliability to reach a greater level of consensus. However, a recent study comparing ordinal and continuous scales in the evaluation of dental arch relationship in a cleft patient population did not find any significant difference between the two types of scales.[42] Finally, it is possible that the expert panel had difficulty differentiating malocclusion severity from treatment difficulty during the rating session. This may have acted as a confounder resulting in decreased inter- and intra-rater reliabilities. A study by Rowe concluded that malocclusion severity and treatment difficulty were related but distinct concepts, and that various components of malocclusion were not reliable predictors of treatment difficulty.[43] More studies are needed to compare the value of ordinal and continuous scales in treatment need index research.

A second limitation was that, for logistical reasons, cases were not randomized among panelists during the expert rating sessions. Cases were also not randomized during the HLD scoring or PAR scoring. This led to some degree of autocorrelation. Autocorrelation was removed and scores were order-adjusted prior to analysis.

Third, the reported state HLD decisions rely on index guidelines only and do not take into account the ability of a consultant to override the HLD decision. The predicted HLD decisions may therefore not reflect the rates of approval and denial that occur in real clinical practice.

Finally, the sample in our study did not consist of any patients with cleft lip and palate, craniofacial anomalies, or severe traumatic deviations, leaving out an important subset of patients with handicapping malocclusions. These patients often utilize different clinical and administrative avenues to seek orthodontic care and therefore rarely present to regular orthodontic offices in Washington state. It is also important to note that Medicaid agencies often have a different process for prior-authorization for these conditions. This may be the reason why these conditions are absent from some state-modified HLD indices. While this may explain the exclusion of these conditions from some HLD indices, it does not explain why in states like MD, MT, NM, TX, and WY, one or more of these conditions are assigned point values that may or may not be equal to the overall point threshold for approval.

More validation studies are needed to validate the orthodontic treatment need indices in the United States. While this study focused on state-modified HLD indices only, these only comprise 42% of all state-modified indices. Many states have indices other than an HLD index which have yet to undergo reliability and validity testing.[9] Future studies on establishing a reliable method of measuring malocclusion severity by an expert orthodontic panel are also needed. While the PAR index is an established, validated, and reliable instrument for measuring malocclusion severity,[1,39] it is not practical in the context of orthodontic expert panels. Ordinal and continuous scales, like the visual analog scale, offer a simple, more time-efficient potential means of evaluating malocclusion severity, but a dearth of comparative studies makes their use unreliable in orthodontic index validation studies.

CONCLUSION

The results of this study suggest there is a lack of consensus among state Medicaid agencies on what constitutes a medically necessary handicapping malocclusion. A person's eligibility of Medicaid-funded orthodontic care appears inconsistent among the 21 states evaluated in this study. Continued modification of HLD indices without supporting validation studies may perpetuate inequities in access to orthodontic care within the Medicaid-eligible orthodontic population in the United States.

AIM 2: Implementation of the HLD index in Washington State

OBJECTIVE

The objective of this study was to conduct a retrospective analysis of the application of Washington State's HLD modified index (HLD(WA)). Providers' determination of medical necessity were compared to the Washington State Medicaid Agency's treatment need decision for all patients screened consecutively at five Seattle-area orthodontic clinics between April and June 2018.

METHODS

Sample

Following University of Washington IRB approval (STUDY00005449), five orthodontic clinics were recruited to participate in this study: one academic clinic, one community health clinic, and three private orthodontic practices. All were located throughout the greater Seattle area. The Washington State HLD index, HLD(WA), was used by all clinics to obtain prior-authorization for orthodontic treatment from Washington Medicaid. All HLD(WA) prior-authorization forms (Appendix 7) submitted consecutively to the Washington Medicaid program on behalf of Medicaid-eligible orthodontic patients between April and June 2018 were collected and de-identified. HLD(WA) prior-authorization forms were matched to a treatment decision letter from the Washington Medicaid program. Inclusion criteria were 1) all consecutive Medicaid-eligible patients whose prior-authorizations were submitted for comprehensive orthodontic treatment to the Washington Medicaid program between April 1 and June 30, 2018 at one of the selected 5 orthodontic clinics, 2) presence of a submitted HLD(WA) prior authorization form, 3) presence of a Washington Medicaid decision letter. Exclusion criteria include 1) screened patients who were not deemed ready to begin orthodontic treatment at the time of screening and 2) patients whose records were missing an HLD(WA) prior authorization or a WA DSHS decision letter.

Each patient's HLD(WA) demographic characteristics (age, gender, date of birth), qualifying exceptions, point-contributing conditions, and overall HLD(WA) provider scores were gathered. Treatment decision information was extracted from the Washington Medicaid decision letter. In the event of a denial, the reasons for denial, consultant HLD(WA) score, and the treatment recommendations were also gathered from the state decision letter.

Data Analysis

Descriptive statistics were used to evaluate clinic and patient demographics, summarize HLD(WA) decisions among clinic types, and compare provider and reviewer scores. Chi-square test for independence was used to compare differences in proportions of enrolled patients as well as approved and denied applications among clinic types. A sub-analysis of approved and denied prior-authorizations was conducted to detect decisions based on index guidelines compared to reviewer override. We also compared the proportions of class I, II, and III malocclusions among approved and denied applications using the Chi-square test for independence. Student's t test was used to compare mean provider and reviewer scores among approved and denied applications. The Kruskal-Wallis test was used to compare mean provider and reviewer scores among clinic types. A multiple regression analysis with a Bonferoni correction set at $p=0.0071$ conducted to quantify the relationship of malocclusion characteristics as listed on the HLD(WA) prior-authorization with an approval outcome by Washington state. Finally, a qualitative descriptive analysis was performed to survey the recommendations made by the Washington DSHS in the event of prior-authorization denial.

RESULTS

Demographics

Clinic types: 1 academic clinic, 1 community health clinic (6 locations), and 3 independent private practices were enrolled in the study. Because the community health center screened a greater number of patients in the 3-month study period, every third consecutive application was enrolled in the study. The community clinic had the largest number of submitting providers (19) followed by the academic clinic (17). During the study period, the prior-authorizations of 479 patients were submitted for agency review and subsequently collected. 10 were excluded due to missing records and 51 were excluded for being submitted outside the study period. 418 prior authorizations were therefore included in the study. There was a statistically significant difference in the number of patients allocated per clinic types for comprehensive treatment, chisq(3) = 126.39, p<.001 (Table 14).

Table 14: Clinic type and site demographics

	# of providers n (%)	*# of patients n (%)*
Academic Clinic	**17 (42.5%)**	**58 (14.1%)**
Clinic A	17	58 (13.9%)
Community Clinics	**19 (47.5%)**	**242 (57.7%)**
Clinic A	3	38 (8.8%)
Clinic B	4	33 (7.9%)
Clinic C	1	6 (1.4%)
Clinic D	4	94 (22.5%)
Clinic E	3	58 (13.9%)
Clinic F	4	13 (3.1%)
Private Clinics	**4 (10.0%)**	**118 (28.2%)**
Clinic A	1	51 (12.2%)
Clinic B	1	14 (3.3%)
Clinic C	2	53 (12.7%)
Total	**40**	**418**
P value (community vs academic vs private)		*<0.001*

P value = chi-squared P value

Patient demographics: Patient demographic characteristics were gathered from the HLD(WA) indices and are presented in Table 15. The mean age was 14.3±5.5 years. The majority of patients (58.8%) had an Angle's Class I malocclusion. Class II and III malocclusions constituted 31.6% and 9.6% of the sample, respectively. 42.6% of prior-authorizations had at least one qualifying exception. Impinging overbite was the most common qualifying exception among patients submitted for comprehensive treatment (22.9%). Mean overjet and overbite were 4.3±2.6mm and 3.7±2.2mm, respectively. Aside from overjet and overbite, anterior crowding was the most prevalent point-contributing condition (58.4%), followed by ectopia (30.4%), and posterior crossbite (15.8%). It should be noted that HLD(WA) guidelines indicate that anterior ectopia and anterior crowding are mutually exclusive and only the most severe condition should be counted if both are present.

Table 15: Demographics for Medicaid-eligible patients (n=418)

	n	*%*
Demographics		
Age in years (SD)	14.3 (5.5)	
Sex		
Male	164	39.2
Female	253	60.5
N/A	1	0.2
Angle Classification		
Class I	244	*58.8*
Class II	131	*31.6*
Class III	40	*9.6*
N/A	3	*0.7*
HLD(WA) components		
Qualifying Exceptions	178	42.6
Impinging overbite	96	*22.9*
Anterior crossbite	38	*9.1*
Overjet >9mm or > -3.5mm	28	*6.7*
Skeletal Cl III	16	*3.8*
Point-contributing Conditions		
Overjet in mm (SD)	4.3 (2.6)	-
Overbite in mm (SD)	3.7 (2.2)	-
Md Protrusion	21	*5.0*
Anterior open bite	29	*6.9*
Ectopia	127	*30.4*
Anterior crowding	244	*58.4*
Posterior crossbite	66	*15.8*

HLD(WA) decisions among clinic types

A summary of HLD(WA) decisions is presented in Table 16. A mean of 35.4% of submitted prior-authorizations were approved for comprehensive treatment across all clinic types, and 53.0% were denied (Figure 6). The Washington Medicaid program was unable to process the remaining 59 applications as a result of missing records or incomplete HLD(WA) indices. There was no statistical difference between the rate of approval and denial treatment decisions per clinic type, chisq(4)=0.582, p=0.7473.

While overall review rates were not different among clinic type, an analysis of application decisions made based on index guidelines compared to reviewer override revealed statistically significant differences among clinic types for both approved (chisq(2)=10.337, p<.01) and denied (chisq(2)=17.363, p<.001) applications (Table 16). Specifically, only 47.6% of academic clinic approvals were based on index guidelines at the academic clinic, compared to 81.6% and 72.5% at the community and private clinics, respectively (Figure 7a). Denial rates followed a similar trend. While the total denial rates were similar among all clinic types (Figure 6), the community clinic had a lower denial rate based on index guidelines (70.5%) and a higher denial rate based reviewer override (29.5%) than both academic and private clinics (Figure 7b).

Table 16: HLD(WA) decisions per clinic type according to index guidelines & reviewer override

	Total n (%)	*Academic Clinic n(%)*	*Community Clinic n(%)*	*Private Clinic n(%)*	*P value*
Total					*0.7473*
Approved	*148 (35.4%)*	21 (36.2%)	87 (36.0%)	40 (33.9%)	
Denied	*219 (52.4%)*	30 (51.7%)	122 (50.4%)	67 (56.8%)	
Unable to process	*51 (12%)*	7 (12%)	33 (14%)	11 (9%)	
Approved by:					<.01
Index guidelines (≥25pts; >0 QE)	110 (74.3%)	10 (47.6%)	71 (81.6%)	29 (72.5%)	
Reviewer override (<25pts; 0 QE)	38 (25.7%)	11 (52.4%)	16 (18.4%)	11 (27.5%)	
Denied by:					<.001
Index guidelines (<25pts; 0 QE)	171 (78.1%)	27 (90.0%)	86 (70.5%)	58 (86.6%)	
Reviewer override (>25pts; >0 QE)	48 (21.9%)	3 (10.0%)	36 (29.5%)	9 (13.4%)	

P value: Chi-squared P value

Figure 6: Approval and denial rates among clinic types; ns = non-significant

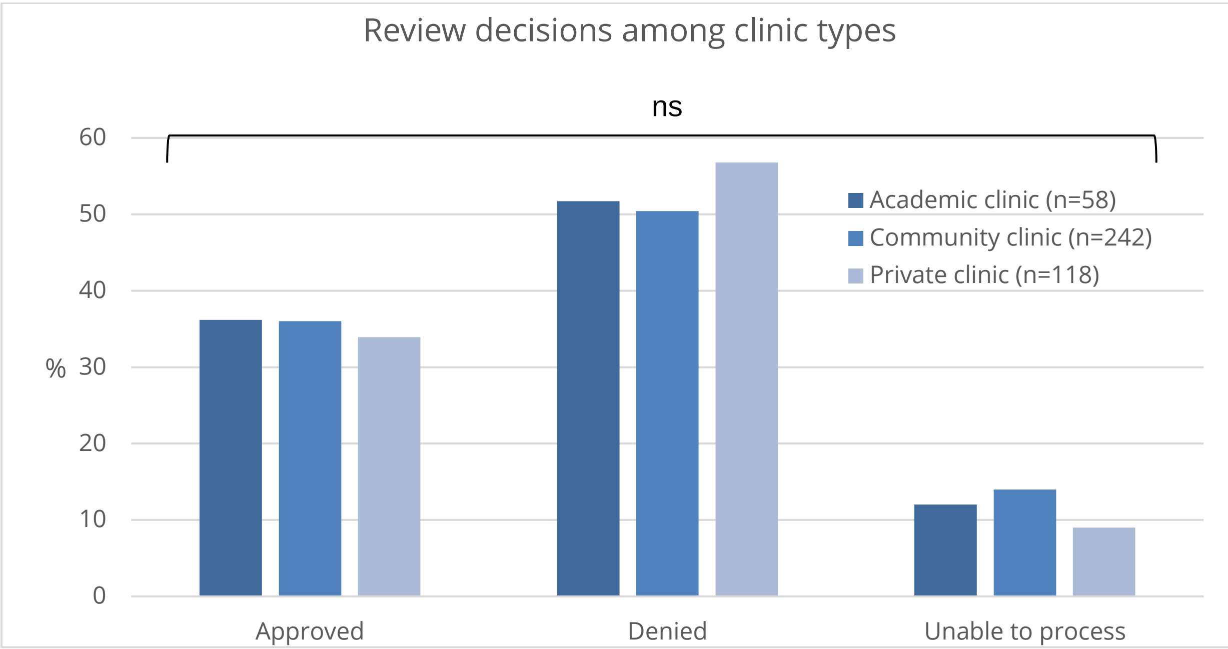


Figure 7: Approvals and Denials by index guidelines and reviewer override per clinic type

Figure 7a

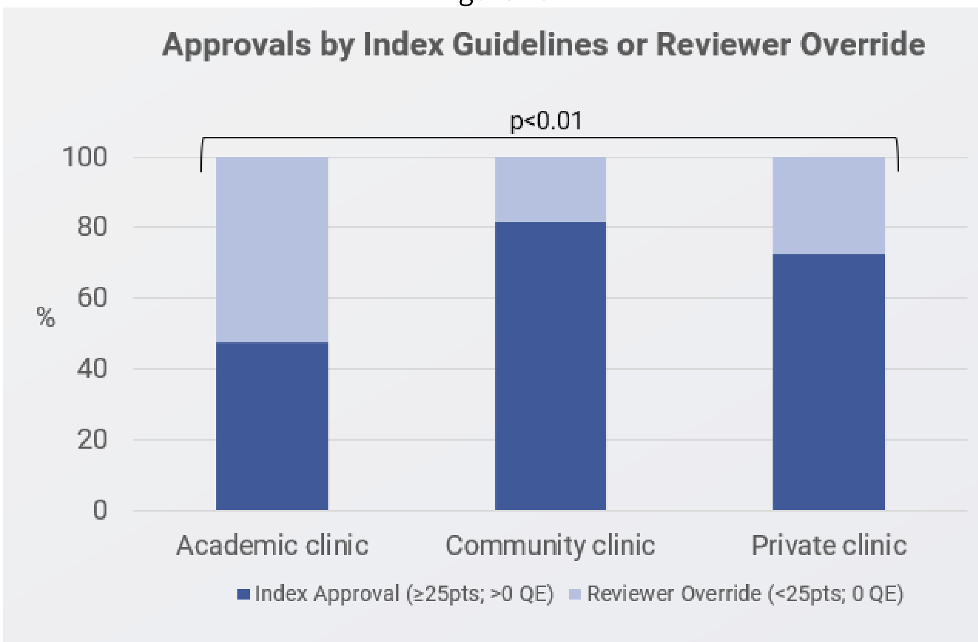
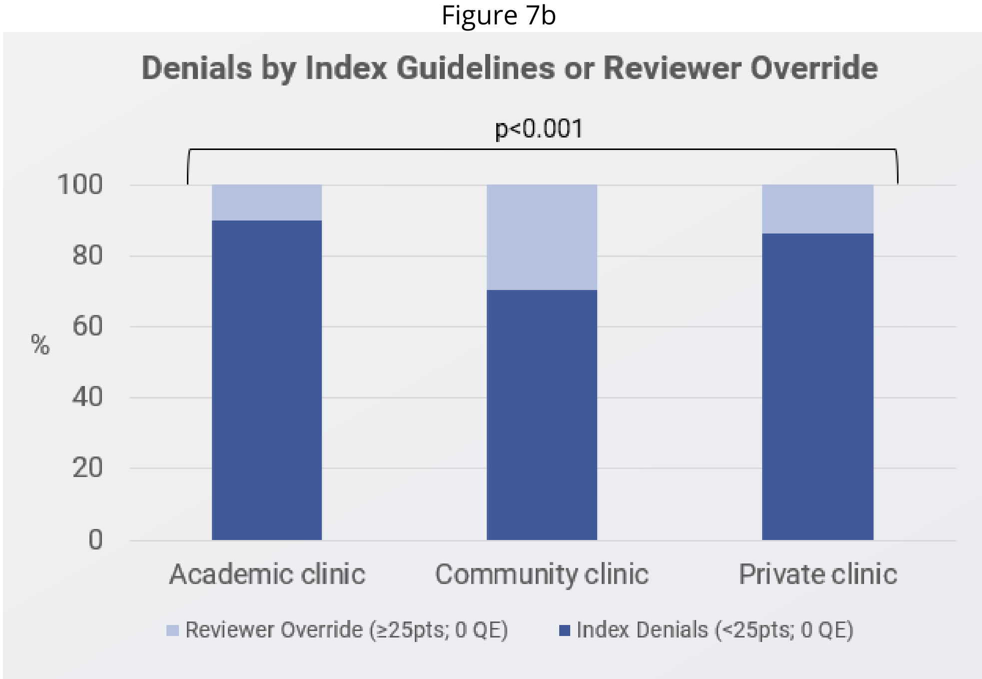


Figure 7b



HLD Score Analysis

Provider Scores: A summary of mean provider scores is presented in Table 17. The provider scores of approved applications (18.1±7.7) were significantly greater than those of denied applications (14.8±6.3), $t(365)=4.42$, $p<.0001$. Provider scores of approved applications were not significantly different whether approvals were based on index guidelines or reviewer override, $t(128)=1.898$, $p=.0600$. However, there was a statistically significant difference in provider scores for applications denied based on index guidelines or reviewer override, $t(71)=-4.689$, $p<.0001$.

When approvals and denials were broken into applications approved based on meeting the HLD(WA) point threshold or having a qualifying exception (Table 18), patients approved on the basis of meeting the point threshold had a significantly greater provider score (19.5±6.8) than patients approved on the basis of qualifying exceptions (17.3±8.2), $t(107)=1.69$, $p<.05$. This was not true for denied applications. The mean provider score for all approved applications was below the HLD(WA) point threshold of 25.

Reviewer scores: Since reviewer scores are only provided in the event of a denial, only denials were included in the analysis. The mean reviewer scores (11.9±5.7) among denied applications was on average 2.9 points lower than the mean provider scores (14.8±6.3), $t(436)=1.97$, $p<.0001$ (Table 17). 42.0% (n=92) of denied applications had equal reviewer and provider scores (Figure 8 & 9). The majority (53%; n=116) of applications had greater provider than reviewer scores, and only 5% had a greater reviewer than provider scores. Of those applications which received a lower reviewer than provider scores, 62.1% were within 5 points and 86.2% were within 10 points of the provider score (Figure 9). It is unclear why one application was denied despite receiving reviewer and provider scores which met index approval guidelines (Figure 8).

The difference between reviewer and provider scores was statistically significant regardless of whether an application was denied for failing to meet index guideline or by reviewer override, $p<.001$ (Table 17). There was a greater provider to reviewer point difference (5.8pts) for applications denied by reviewer override than for applications denied for failing to meet index guidelines (2.0pts).

Of all denied applications, 38 applications had at least one qualifying exception (Table 18). Both the provider and reviewer scores were lower than the required 25 point threshold in 92.1% (n=35) of denied cases citing at least one qualifying exception. 3 had a provider score greater than the index threshold which was later adjusted to below the 25 point threshold by the reviewer. While patients denied with and without qualifying exceptions did not have significantly different provider scores, the mean reviewer scores were statistically significantly lower, $p<.001$. The provider to reviewer score difference was greater for denied applications with qualifying exceptions (mean difference=5.0) compared to applications with no qualifying exceptions (mean difference=2.4).

Both provider ($p<.05$) and reviewer scores ($p=.001$) were significantly different among clinic types for patients denied orthodontic treatment (Table 19 and Figure 10). The mean reviewer scores were lower than the provider scores in all three clinic types. A post-hoc analysis revealed that both academic ($p=.005$) and private clinic ($p=.021$) reviewer scores were significantly lower than community clinic reviewer scores.

Table 17: Mean provider and reviewer scores by approval and denial status

	Total			*Approved by:*			*Denied by:*		
	Approved (n=148)	*Denied (n=219)*	*P value*	*Index (n=110)*	*Reviewer override (n=38)*	*P value*	*Index (n=164)*	*Reviewer override (n=55)*	*P value*
Mean provider score (SD)	18.1 (7.7)	14.8 (6.3)	*<.0001*	18.6 (8.6)	16.5 (4.2)	*.0600*	13.5 (5.2)	18.7 (7.8)	*<.0001*
Mean reviewer score (SD)	-	11.9 (5.7)	-	-	-	-	11.5 (5.0)	12.9 (7.4)	*.2085*
Score difference	-	2.9	-	-	-	-	2.0	5.8	-
P value	-	*<.0001*	-	-	-	-	*<.001*	*<.0001*	-

P value = Student t test one-tail P value

Table 18: Mean provider and reviewer scores by qualifying exception or point-contributing conditions

	Provider Score $\bar{x}$ (SD)	*Reviewer Score $\bar{x}$ (SD)*	*P Value*
Approved with:			
≥1 qualifying exception (n=48)	17.3 (8.2)	-	-
No qualifying exception (n=100)	19.5 (6.8)	-	-
P Value	*.0466*	-	-
Denied with:			
≥1 qualifying exception (n=38)	15.0 (6.2)	10.0 (6.7)	*<.001*
No qualifying exception (n=181)	14.7 (6.4)	12.3 (5.5)	*<.0001*
P Value	*.3992*	-	-

P value = Student t test one-tail P value

Table 19: Mean provider and reviewer scores among clinic types

	Academic clinic (n=30)		*Community clinic (n=122)*		*Private clinic (n=67)*		*Total (n=219)*		
	$\bar{x}$	*SD*	*$\bar{x}$*	*SD*	*$\bar{x}$*	*SD*	*$\bar{x}$*	*SD*	*P Value*
Mean provider score	13.3	6.7	15.8	5.8	13.7	6.8	14.8	6.3	0.037
Mean reviewer score	9.5	4.6	12.9	5.9	11.0	5.4	11.9	5.7	0.001
Score difference	3.8	4	2.8	4.2	2.7	5.2	2.9	4.6	

P value = Kruskal-Wallis P value

Figure 8: Provider and reviewer scores among patients denied comprehensive orthodontic treatment (n=219)

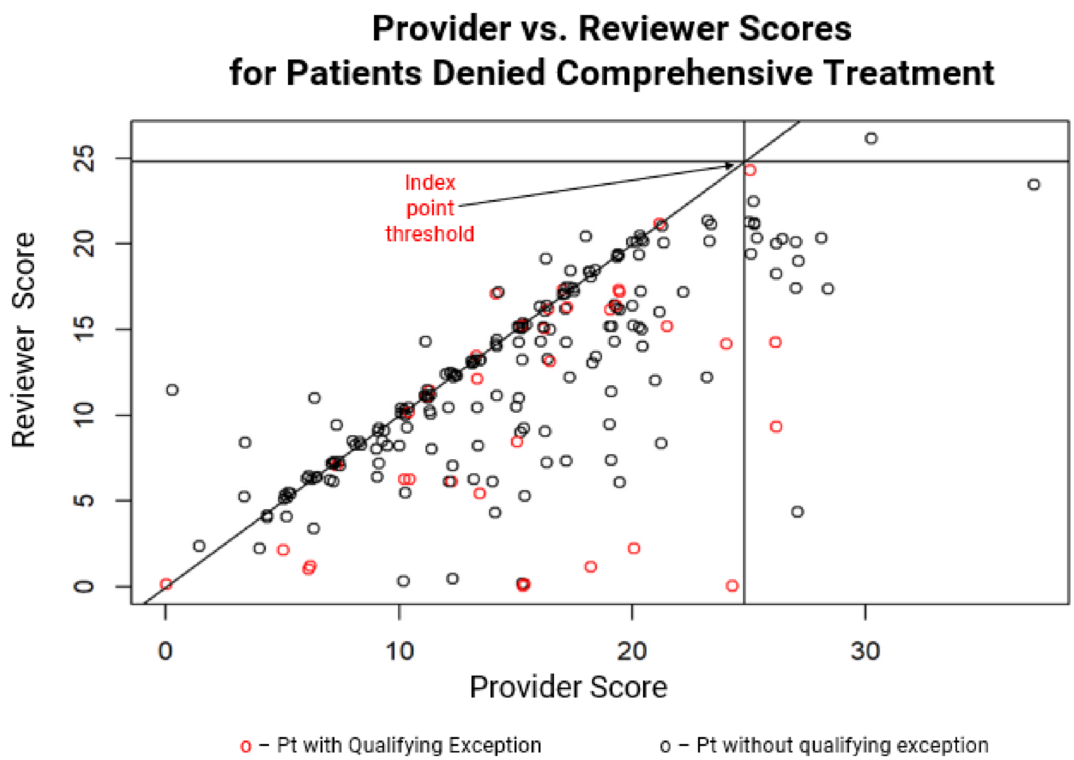


Figure 9: Distribution of provider and reviewer score differences for patients denied comprehensive orthodontic treatment (n=219)

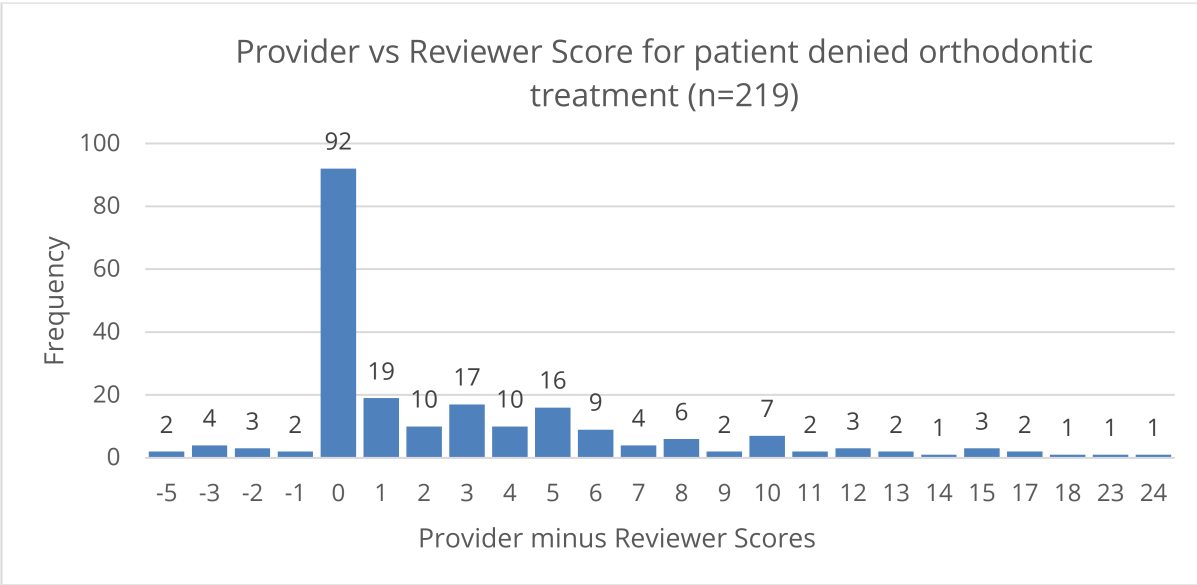
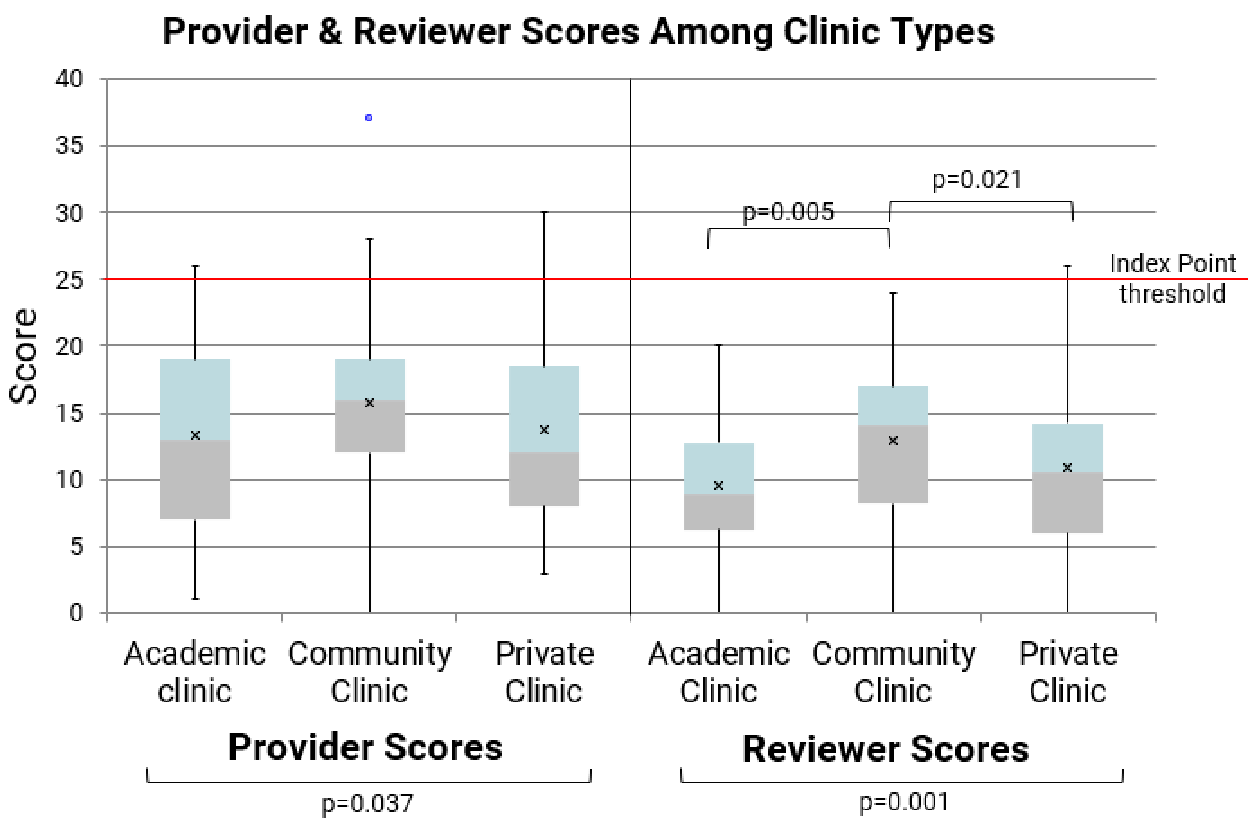


Figure 10: Provider and reviewer scores among clinic types for patient denied orthodontic treatment (n=219)



Influence of malocclusion characteristics on HLD treatment decision:

The distribution of Class I, II, & III malocclusions among patients who were approved and denied comprehensive treatment is presented in Table 20. There was a significant difference in Angle classification between approved and denied applications, chisq(2)=23.5, p<.0001. Among the patients who were approved for orthodontic treatment, there was an equal proportion of Class I and II malocclusions (43.2% & 43.9%, respectively). Contrarily, more patients who were denied treatment tended to have a Class I malocclusion (68%). Class III malocclusions were only slightly more prevalent among patients who were approved (12.8%) compared to those who were patients (7.8%).

Table 20: Angle classification among Medicaid recipients in Washington State (n=418)

	Approved		*Denied*		*P value*
	n	*%*	*n*	*%*	
Angle's Classification					<.0001
Class I	64	43.2	149	68.0	
Class II	65	43.9	51	23.3	
Class III	19	12.8	17	7.8	
N/A	-	-	2	0.9	

P value: Chi-squared *P* value

A multiple regression analysis was also used to quantify the relationship of point-contributing conditions with the outcome of Washington State approval (Table 21). Qualifying exceptions were not included in this analysis since their presence should lead to an approval decision. Among the point-contributing conditions, overbite (p<.001), mandibular protrusion (p=.001), anterior open bite (p<.001), and posterior unilateral crossbite (p=.002) were all significantly related with an approval outcome in Washington state.

Table 21: Odds Ratios of point-contributing conditions using HLD(WA)

Predictors	*Odds Ratios*	*CI*	*P value*
Overjet	1.13	1.02-1.25	0.024
Overbite	1.29	1.14-1.48	<0.001
Md protrusion	1.33	1.15-1.64	0.001
Ant. Open Bite	1.23	1.11-1.39	<0.001
Ant. Ectopic Teeth	0.95	0.89-1.02	0.166
Ant. Crowding	1.00	0.95-1.06	0.876
Posterior unilateral crossbite	1.21	1.08-1.37	0.002

with Bonferroni correction

Washington HCA clinical recommendations

DSHS recommendations for future treatment were tabulated (Table 22). These were divided into primary and secondary reasons, with the former being the main explanation for the agency's treatment decision, and the latter being the case-specific reason and recommendation. Two primary reasons for denial were identified: 1) failure to establish medical necessity and 2) request for additional information not answered. Failure to establish medical necessity was the most commonly cited reason for denial (99.5%). Once a patient failed to meet medical necessity, the agency didn't provide an explanation 77.2% of the time, reported that the photos did not validate the reported condition 20.1% of the time, determined that oral hygiene was too poor for orthodontic treatment in 1.8% of cases. The most common recommendation for future treatment was no treatment (61.2%), followed by interceptive treatment (7.8%), and limited treatment (6.0%).

Table 22: Reasons for denial and agency recommendation for future treatment

Agency Denial Response	*n*	*%*
Failure to establish medical necessity	**218**	99.5
No explanation, agency recommends:	169	77.2
No treatment	134	61.2
Extractions and allow for spontaneous eruption	1	0.5
TMJ workup with a splint, if TMD is suspected	3	1.4
Interceptive Treatment	16	7.3
Limited Treatment	12	5.5
Nightguard to control effects of night grinding	2	0.9
Denied until OS can determine medical necessity	1	0.5
Photos do not validate reported condition, agency recommends:	44	20.1
No treatment	42	19.2
Remove braces & deliver retainers	-	-
Interceptive Treatment	1	0.5
Limited Treatment	1	0.5
Poor oral hygiene, agency recommends:	4	1.8
Resubmit in 6 months when oral hygiene has improved	4	1.8
Primary teeth still present, agency recommends:	1	0.5
No treatment	1	0.5
Request for additional information not answered	**1**	0.5
Request a reconsideration of this decision	1	0.5
Total	*219*	100

DISCUSSION

Access to dental care is a widespread problem for Medicaid-eligible patients.[44] Factors known to influence a patient's access to dental care include geographic barriers,[45,46] declining provider participation in Medicaid programs,[47] and racial and ethnic factors.[48,49] The current study shed some light on an additional barrier which has yet to be reported in the orthodontic literature: inter-clinic disparities.

Our study found that overall approval and denial rates for comprehensive orthodontic treatment were 35.4% and 52.4% respectively. The Medicaid program could not process the remaining 12.2% of applications due to clerical reasons. These results closely mirror those of a 2005 study which found a 35% approval rate among Medicaid patients in Washington state.[26] Meanwhile a 2017 study found a 54% approval rate among Medicaid applications in Illinois.[27] Because state Medicaid programs regularly update and modify the indices' qualifying conditions and point thresholds, a comparison of HLD approval and denial rates across states and through time can be challenging. In Washington state alone, the HLD index has experienced considerable changes to its component criteria and point-threshold over the past several years. The HLD(WA) used at the time of the 2005 study had a point-threshold of 30 and included a "labio-lingual spread" component.[26] The "labio-lingual spread" component is no longer part of the 2019 HLD(WA) and the point threshold has since been lowered to 25. Similarly, the HLD index used by the state of Illinois used different qualifying exceptions, a point threshold of 28, and a "labio-lingual spread" component, making it qualitatively different from the HLD(WA) used in the present study.[27] Other studies also reported approval and denial rates, but because patient records were not reviewed by a Medicaid agency consultant, but rather by orthodontists as part of an experiment, a comparison was not possible.[15,17,23] For all reasons stated above, a comparison of approval and denial rates is difficult to establish using the current body of work. More studies focused on evaluating Medicaid agency reviewed cases are needed to accurately assess and compare the implementation of orthodontic indices for publicly-funded orthodontic care.

While at first glance, approval and denial decisions seem to be similar among academic, community, and private clinics, a more in-depth analysis shows that clinic type may have an effect patients' application approval outcome. Notably, this study found that a significantly greater proportion of academic clinic patients were getting approved when their HLD(WA) applications did not meet index guidelines for approval. This finding could be the result of a couple of factors. For one, the academic clinic also had a highly heterogeneous pool of providers. Since HLD(WA) application indices at the academic clinic are completed by students who have less orthodontic experience than either the private and community clinics, where orthodontists complete the applications, it is possible that systematic errors are being made which Medicaid program consultants subsequently correct. While reviewer score adjustments are not provided in the event of an approval decision, our data among denied applications suggests that the provider to reviewer score difference is greatest in the academic clinic, thereby supporting this explanation. Another explanation could be that Medicaid program consultants give preferential treatment to the academic clinic compared to other clinic types. If this were the case, we would expect total approval rates to be greater for the academic clinic compared to other clinic types. This was not the case.

Another significant finding of this study was that the community clinic had a greater percentage of denials based on reviewer override. This may be the result of providers at the community clinic inflating the severity of their patient's malocclusions in an effort to increase the probability of an approval outcome. This explanation is supported by the findings that reviewer scores were consistently and significantly lower than provider scores, and that among the 38 applications denied prior-authorization in all clinic types, 29 (76.3%) originated from the community clinic. The most cited qualifying exceptions were deep impinging overbite (n=12) and anterior crossbite of individual teeth (n=17). In addition, 3 of the 38 denied applications also had a score equal to or greater than the index point threshold, but all 3 scores were adjusted by the reviewer to a score lower than the index threshold. These observations reinforce the hypothesis that in all denied cases citing at least one qualifying exception, the reviewer did not agree with the provider that a qualifying exception was in fact present. Another possibility for explaining the greater number of reviewer override-based denials in the community clinic is that the consultant may apply a more stringent review of applications emanating from these clinics. If this were the case, however, we would expect to see a higher overall denial rate in the community clinic. This study did not support this observation. In addition, these findings could not be compared to previously published results since, to our knowledge, no study has previously compared Medicaid orthodontic decisions rates among clinic types.

Our analysis of the provider and reviewer scores reveals that in 53% of denials, reviewer scores were lower than the submitted provider scores. This could be a result of providers inflating their patient's scores to increase the chance of an approval outcome, inadequate training and calibration among providers using the HLD(WA), inadvertent errors in completing the HLD(WA) index, ambiguity in the HLD(WA) component criteria, or professional provider-reviewer disagreement of what constitutes any specific condition as listed and explained on the index. The common factor in all these explanations is the fact that malocclusion severity is defined by the amount of morphological deviation from an idealized norm and is particularly difficult to quantify. As a result, treatment need indices often fall short of having the ideal internal validity and inter-rater reliability.[15,17,31] This problem is further complicated by the fact that indices include nebulous language which may create some confusion among providers, but also leaves room for interpretation for both providers and consultants alike.[17,31] Medicaid programs, including the Washington Medicaid agency, compensate for these complications by progressively changing the index over time and by allowing their consultants to override a provider's assessment, as demonstrated in this study by the differences in provider and reviewer score.[14,27,50]

The Washington Medicaid agency has in the past attempted to improve its index. For example, it added a criterion for bilateral crossbite, removed the labio-lingual spread component, and added a qualifying exception not present in other state-modified HLD indices ("negative overjet due to skeletal Class III"). It has also lowered its point-threshold, which studies have shown increases the instrument's sensitivity[15,23] and approval rate.[27] While these changes may lead to an incremental improvement of the Washington-modifed HLD index, they are slow to come and lack supporting reliability and validity studies. This is a concern since Medicaid legislation requires that limited funds be allocated to patients in greatest need of medically-necessary orthodontic care.[3] To overcome these concerns, state Medicaid agencies rely on consultants to be the gatekeepers for publicly-funded orthodontic care and have put in place appeal processes should a patient disagree with a consultant's decision.[35] While this administrative process provides a means to make up for the HLD(WA) index' deficiencies, it is not a long-term solution. Instead, a modification process driven by index validation studies is more likely to provide uniformity and equity among Medicaid recipients.[1]

Several recommendations can be made based on the findings of this study to improve the HLD(WA) index and Medicaid patients' access to equitable orthodontic care in Washington State. One recommendation is that the agency blind its consultants to the submitting provider and clinic type. Such a measure would eliminate any clinic-based biases and ensure that consultants are reviewing submitted HLD(WA) prior-authorizations for their adherence to state-approved guidelines. We predict this would also minimize the rate of consultant overrides, which in this study changed the expected HLD(WA) decision in 20.5% of applications. Another recommendation is for state Medicaid programs to rely on rigorous validation/reliability studies to guide any modification of the index in use. Once an index has been validated, only the point threshold should be modified to meet state budgetary demands, as modifying the index itself invalidates any results from previous validation studies. As previous investigators have reported, Medicaid agencies should also seek to clarify the definitions for any index component criteria.[17] For example, despite widely reported confusion surrounding the HLD index' definition of "ectopia",[17,31] the HLD(WA) falls short of mentioning "impactions" despite clarifying that ectopic teeth include "teeth in the maxillary sinus and, in the ascending ramus of the mandible and other such situations, when teeth develop in other locations, rather than in the dental arches."[50] This level of ambiguity only leads providers to inflate or leave out potential qualifying conditions and results in greater reliance on consultant's clinical judgement for approval. Another concern with the current reviewing process is the possibility for reviewer personal biases to affect application decision outcomes. Applying a system in which two consultants review each application could foster stricter adherence to index guidelines and minimize reviewer biases. Finally, a process of provider and reviewer calibration may be necessary to limit biases. Provider calibration could improve competence with regards to the index, potentially limiting the frequency of errors or misinterpretation. Reviewer calibration could ensure uniformity and limit personal biases among reviewers.

Future studies are needed to validate state-modified indices throughout the United States. In addition, more studies are needed to evaluate the implementation of the HLD indices among Medicaid agencies. While the present study focused on evaluating the review process for Medicaid-funded comprehensive orthodontic treatment prior-authorization, it did not evaluate either interceptive or limited treatment modalities. The publicly-funded limited and interceptive treatment prior authorization process often does not rely on any index or set criteria, and is therefore more dependent on the consultants' clinical judgement. Given that there is evidence that limited treatment in the mixed dentition is efficacious in decreasing malocclusion severity as measured by the HLD index,[26,51] research on the implementation of these treatment types could only increase access to orthodontic care.

Limitations and Future Studies

This study had several limitations. A significant limitation of the present study is that HLD indices may not be representative of patients' true malocclusions due to provider biases and index proficiency. No data could be collected on the providers and their levels of experience with the HLD(WA) is therefore unknown. Secondly, there was an uneven distribution of patients among the different clinic types, resulting in over-representation of the community clinic. Because there were many more consecutive patients enrolled in the community clinic, we enrolled every third patient into the study. Third, there was great heterogeneity among providers completing the HLD(WA) indices. The academic and community clinics tended to have more providers than private clinics, and the providers at the academic clinic likely had less orthodontic experience than either the community or private clinics. These factors could all act to decrease the sensitivity of the HLD(WA) index as an instrument for identifying orthodontic need. Another limitation of the study is provider treatment bias. It is likely that providers inflate malocclusion severity to boost their patient's chances of being approved for orthodontic treatment, thereby artificially increasing submitted HLD scores. While state Medicaid program reviewers likely adjust HLD(WA) scores and deny patients whose HLD(WA) scores were artificially inflated, the overall sensitivity of the HLD(WA) as a screening instrument is likely decreased by these confounders.

CONCLUSION

When evaluating access to orthodontic care among a Medicaid population, public health officials, providers, and state Medicaid program reviewers alike should consider that clinic type may have an effect on patients' ability to access publicly funded orthodontic care. This may create inequities within the Medicaid-insured orthodontic patient population, which when compounded with known barriers to care, could create significant access to care problems for this vulnerable population. These inequities can be minimized by ensuring that providers more closely adhere to index guidelines through calibration training, by blinding state-contracted reviewers to the submitting clinic sites, and through more index validation studies.

GLOSSARY OF TERMS

- **Gold standard:** In the context of orthodontic treatment need indices, the gold standard is considered to be the mean opinion of a panel of orthodontist or qualified individuals regarding a patient's orthodontic treatment need. The gold standard is often measured on a likert scale corresponding to nominal treatment need criteria (ex: no treatment needed, treatment advisable, treatment recommended, etc…)

- **Negative predictive value**: The probability that subjects testing negative for handicapping malocclusion test truly do not have a handicapping malocclusion.

- **Positive predictive value:** The probability that subjects testing positive for handicapping malocclusion truly have a handicapping malocclusion.

- **Reliability:** The ability of an index to consistently and repeatedly produce the same results with the same or different raters (Aaronson et al. 2002).

- **Specificity:** The ability of a test to identify healthy persons as not having the condition of interest (i.e. the true negative rate) (Centers for Disease Control and Prevention (CDC) 2015).

- **Sensitivity:** The ability of a test to identify true cases or the proportion of people with a condition identified as such (i.e. the true positive rate) (Centers for Disease Control and Prevention (CDC) 2015).

- **Treatment need:** Need is defined as medically necessary handicapping malocclusion (Parker 1998). In the context of orthodontic treatment need indices, a person with treatment need is defined as a person with an overall score greater than the cut-off score, or the existence of a qualifying exception.

- **Validity:** The validity of an orthodontic treatment need index is defined as the ability of the index to measure what it seeks to measure, i.e. treatment need, and is often measured by comparing its results to the opinions of an expert orthodontic panel, or "gold standard" (Brook and Shaw 1989; Beglin et al. 2001; Aaronson et al. 2002; Arruda 2008).

REFERENCES

1. Shaw WC, Richmond S, O'Brien KD. The use of occlusal indices: A European perspective. *Am. J. Orthod. Dentofac. Orthop.* 1995;107(1):1–10.

2. Health Resources & Services Administration. Early Periodic Screening, Diagnosis, and Treatment | Maternal and Child Health Bureau. 2018. Available at: https://mchb.hrsa.gov/maternal-child-health-initiatives/mchb-programs/early-periodic-screening-diagnosis-and-treatment. Accessed January 12, 2020.

3. Anon. Early and Periodic Screening, Diagnostic, and Treatment | Medicaid. Available at: https://www.medicaid.gov/medicaid/benefits/early-and-periodic-screening-diagnostic-and-treatment/index.html. Accessed January 12, 2020.

4. Draker HL. Handicapping labio-lingual deviations: A proposed index for public health purposes. *Am. J. Orthod.* 1960;46(4):295–305.

5. Draker HL. American association of orthodontists approval of the assessment record form and the definition of handicapping malocclusion. *J. Am. Dent. Assoc.* 1967.

6. Summers CJ. The occlusal index: A system for identifying and scoring occlusal disorders. *Am. J. Orthod.* 1971;59(6):552–67.

7. Grainger RM. Orthodontic treatment priority index. *Vital Health Stat. 2.* 1967;(25):1–49. Available at: http://www.ncbi.nlm.nih.gov/pubmed/5300095. Accessed January 12, 2020.

8. Salzmann JA. Handicapping malocclusion assessment to establish treatment priority. *Am. J. Orthod.* 1968;54(10):749–65.

9. Minick G, Tilliss T, Shellhart WC, et al. Comparison of Orthodontic Medicaid Funding in the United States 2006 to 2015. *Front. Public Heal.* 2017;5. Available at: http://journal.frontiersin.org/article/10.3389/fpubh.2017.00221/full. Accessed January 12, 2020.

10. Cons NC, Jenny J, Kohout FJ, DDS YS, Jotikastira D. Utility of the Dental Aesthetic Index in Industrialized and Developing Countries. *J. Public Health Dent.* 1989;49(3):163–6. Available at: http://doi.wiley.com/10.1111/j.1752-7325.1989.tb02054.x. Accessed January 12, 2020.

11. Brook PH, Shaw WC. The development of an index of orthodontic treatment priority. *Eur. J. Orthod.* 1989;11(3):309–20.

12. Daniels C, Richmond S. The Development of the Index of Complexity, Outcome and Need (ICON): *https://doi.org/10.1093/ortho.27.2.149* 2019.

13. American Association of Orthodontists. No Title. *AAO Bull.* 1990.

14. Parker WS. The HLD (CalMod) index and the index question. *Am. J. Orthod. Dentofac. Orthop.* 1998;114(2):134–41.

15. Younis JW, Vig KWL, Rinchuse DJ, Weyant RJ. A validation study of three indexes of orthodontic treatment need in the United States. *Community Dent. Oral Epidemiol.* 1997;25(5):358–62. Available at: http://doi.wiley.com/10.1111/j.1600-0528.1997.tb00955.x. Accessed January 12, 2020.

16. Beglin FM, Firestone AR, Vig KWL, Beck FM, Kuthy RA, Wade D. A comparison of the reliability and validity of 3 occlusal indexes of orthodontic treatment need. *Am. J. Orthod. Dentofac. Orthop.* 2001;120(3):240–6.

17. Cooke M, Gerbert B, Gansky S, Miller A, Nelson G, Orellana M. Assessment of the validity of HLD (CalMod) in identifying orthodontic treatment need. *Community Dent. Oral Epidemiol.* 2010;38(1):50–7. Available at: http://doi.wiley.com/10.1111/j.1600-0528.2009.00506.x. Accessed January 12, 2020.

18. American Association of Orthodontists. AAO Leads Effort to Establish Consistency on Medically Necessary Orthodontic Care. *AAO Bull.* 2015. Available at: https://www1.aaoinfo.org/aao-leads-effort-standardize-medically-necessary-orthodontic-care-criteria/.

19. American Association of Orthodontists. Medically Necessary Orthodontic Care: AAO Initiative Advances. *AAO Bull.* 2017. Available at: https://www.aaoinfo.org/news/2017/04/medically-necessary-

orthodontic-care-aao-initiative-advances.

20. Proffit WR, Fields HW, Moray LJ. Prevalence of malocclusion and orthodontic treatment need in the United States: estimates from the NHANES III survey. *Int. J. Adult Orthodon. Orthognath. Surg.* 1998;13(2):97–106.

21. Lohr KN. Assessing health status and quality-of-life instruments: Attributes and review criteria. *Qual. Life Res.* 2002;11(3):193–205.

22. Arruda AO. Occlusal indexes as judged by subjective opinions. *Am. J. Orthod. Dentofac. Orthop.* 2008;134(5):671–5.

23. Han H, Davidson WM. A useful insight into 2 occlusal indexes: HLd(Md) and HLD(CalMod). *Am. J. Orthod. Dentofac. Orthop.* 2001;120(3):247–53.

24. Dicker R, Coronado F, Koo D, Parrish R. *Principles of epidemiology in public health practice: an introduction to applied epidemiology and biostatistics. Centers for disease control and prevention.*; 2006. Available at: http://stacks.cdc.gov/view/cdc/6914. Accessed January 12, 2020.

25. Parker WS. A study of 1000 malocclusions selected by the HLD (CalMod) Index. *Am. J. Orthod. Dentofacial Orthop.* 1999;115(4):343–51.

26. Theis JE, Huang GJ, King GJ, Omnell ML. Eligibility for publicly funded orthodontic treatment determined by the handicapping labiolingual deviation index. *Am. J. Orthod. Dentofac. Orthop.* 2005;128(6):708–15.

27. Caplin J, Rozgony A, Farhi M, Viana GC, Al-Grouz M, Belavsky BZ. The implementation of the Handicapping Labio-Lingual Deviation index on access to orthodontic care. *J. Public Health Dent.* 2019:jphd.12352. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/jphd.12352. Accessed April 29, 2020.

28. Bollen A-M. Effects of Malocclusions and Orthodontics on Periodontal Health: Evidence from a Systematic Review. *J. Dent. Educ.* 2008;72(8).

29. Dimberg L, Arnrup K, Bondemark L. The impact of malocclusion on the quality of life among children and adolescents: a systematic review of quantitative studies. *Eur. J. Orthod.* 2015;37(3):238–47. Available at: https://academic.oup.com/ejo/article-lookup/doi/10.1093/ejo/cju046. Accessed January 12, 2020.

30. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* 2016;15(2):155–63.

31. Parker WS. Useful data from application of the HLD (CalMod) Index. *Am. J. Orthod. Dentofacial Orthop.* 2000;117(4):435–7.

32. Anon. Early Periodic Screening, Diagnosis, and Treatment | Maternal and Child Health Bureau. Available at: https://mchb.hrsa.gov/maternal-child-health-initiatives/mchb-programs/early-periodic-screening-diagnosis-and-treatment. Accessed May 3, 2020.

33. Oppenhuizen G. *Medically Necessary Orthodontic Care: Index or Auto-Qualifier; Correlation and Simplicity*.; 2018.

34. Warehouse MD. *Nevada Department of Helath and Human Services Division of Public and Behavioral Helath - Oral Health Program*. Las Vegas, NV

35. Anon. *EPSDT-A Guide for States: Coverage in the Medicaid Benefit for Children and Adolescents Early and Periodic Screening, Diagnostic and Treatment (EPSDT)Medicaid-CHIP-Program-Information/By-Topics/Benefits/Early-and-Periodic-Screening-Diagnostic-and-Treatmen*.; 2014. Available at: https://www.medicaid.gov/medicaid/benefits/early-and-periodic-screening-diagnostic-and-treatment/index.html. Accessed May 4, 2020.

36. Breistein B, Burden DJ. Equity and orthodontic treatment: a study among adolescents in Northern Ireland. *Am. J. Orthod. Dentofacial Orthop.* 1998;113(4):408–13.

37. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. *J. Clin. Nurs.* 2005;14(7):798–804. Available at: http://doi.wiley.com/10.1111/j.1365-2702.2005.01121.x. Accessed

May 3, 2020.

38. Kiekens RMA, Maltha JC, Van 't Hof MA, Kuijpers-Jagtman AM. A measuring system for facial aesthetics in Caucasian adolescents: reproducibility and validity. *Eur. J. Orthod.* 2005;27:579–84. Available at: https://academic.oup.com/ejo/article-abstract/27/6/579/400863. Accessed May 3, 2020.

39. Richmond S, Shaw W, O'Brien K, et al. The development of the PAR Index (Peer Assessment Rating): reliability and validity. *Eur. J. Orthod.* 1992;14(2):125–39. Available at: https://www.ncbi.nlm.nih.gov/pubmed/1582457. Accessed May 3, 2020.

40. Phillips C, Tulloch C, Dann C. Rating of facial attractiveness. *Community Dent. Oral Epidemiol.* 1992;20(4):214–20. Available at: http://doi.wiley.com/10.1111/j.1600-0528.1992.tb01719.x. Accessed May 4, 2020.

41. Schabel BJ, McNamara JA, Franchi L, Baccetti T. Q-sort assessment vs visual analog scale in the evaluation of smile esthetics. *Am. J. Orthod. Dentofac. Orthop.* 2009;135(4 SUPPL.):S61–71.

42. Fowler P, Bellardie H, Shaw B, Eyres P, Semb G, Thompson J. Reliability of a Categorical Scale (GOSLON) and a Continuous Scale (10-cm Visual Analog Scale) for Assessing Dental Arch Relationships Using Conventional Plaster and 3D Digital Orthodontic Study Models of Children With Complete Unilateral Cleft Lip and Palate. *Cleft Palate-Craniofacial J.* 2019;56(1):84–9. Available at: http://journals.sagepub.com/doi/10.1177/1055665618770054. Accessed May 4, 2020.

43. Rowe KGT. The concordance of pre-treatment malocclusion assessment among orthodontic specialty practitioners. 1989.

44. Colby M, Natzke B. Health care utilization among children enrolled in Medicaid and CHIP via Express Lane Eligibility. *Health Serv. Res.* 2015;50(3):642–62.

45. Byck GR, Walton SM, Cooksey JA. Access to Dental Care Services for Medicaid Children: Variations by Urban/Rural Categories in Illinois. *J. Rural Heal.* 2002;18(4):512–20. Available at: http://doi.wiley.com/10.1111/j.1748-0361.2002.tb00918.x. Accessed April 29, 2020.

46. McKernan SC, Kuthy RA, Momany ET, et al. Geographic accessibility and utilization of orthodontic services among Medicaid children and adolescents. *J. Public Health Dent.* 2013;73(1):56–64. Available at: http://doi.wiley.com/10.1111/jphd.12006. Accessed April 29, 2020.

47. Edelstein BL. Examining whether dental therapists constitute a disruptive innovation in US dentistry. *Am. J. Public Health* 2011;101(10):1831–5.

48. Merritt JM, Greenlee G, Bollen AM, Scott JM, Chi DL. Racial disparities in orthodontic service use for Medicaid-enrolled children: An evaluation of the Washington Medicaid program. *Am. J. Orthod. Dentofac. Orthop.* 2016;149(4):516–22.

49. Okunseri C, Pajewski NM, McGinley EL, Hoffmann RG. Racial/Ethnic Disparities in Self-Reported Pediatric Orthodontic Visits in the United States. *J. Public Health Dent.* 2007;67(4):217–23. Available at: http://doi.wiley.com/10.1111/j.1752-7325.2007.00032.x. Accessed April 29, 2020.

50. Anon. Forms and Publications | Washington State Health Care Authority. Available at: https://www.hca.wa.gov/billers-providers-partners/forms-and-publications?combine=&field_topic_tid=15586&field_billers_document_type_value_1=All&sort=filename ASC. Accessed May 3, 2020.

51. King GJ, Spiekerman CF, Greenlee GM, Huang GJ. Randomized Clinical Trial of Interceptive and Comprehensive Orthodontics. *J. Dent. Res.* 2012;91:S59–64.

APPENDICES

APPENDIX 1a: Qualifying exception utilization among state-modified HLD indices, with weight factor

	AK	*AR*	*CA*	*DC*	*DE*	*FL*	*IL*	*MA*	*MD*	*ME*	*MO*	*MT*	*NJ*	*NM*	*NY*	*OK*	*RI*	*SC*	*TX*	*WA*	*WY*	*AAO*
Cleft Palate deformity	QE	QE[1]	QE	QE	QE	QE	QE	QE	15pts (QE)	QE	QE	QE	QE	QE	QE	QE	QE[38]	QE	15pts	QE[1, 38]	QE	QE
Craniofacial anomaly		QE[1]	QE	QE		QE[1]		QE		QE[1]	QE[1]	QE[1]	QE	QE[1]	QE	QE[1]		QE	QE[1]	QE[1]		QE
Impinging overbite	QE[2]		QE[2]	QE[2]	QE[2]	QE[2]	QE[2]	QE[2]		QE[2]	QE[2]	QE[2]	3pts[2]	QE[2]	QE[2]	"X"[2, 21]	QE[2]			QE[2, 3]	QE[2]	QE[2]
Severe traumatic deviation	QE		QE	QE	QE	QE	QE	QE	15pts (QE)	QE	QE	15pts	QE	15pts	QE		15pts		15pts		15pts	QE
Crossbite of ant. teeth	QE[2]		QE[2]	QE	QE[2, 28]	QE[2, 28]	QE[2]			QE[2]	QE[2]	4pts	QE	QE[2, 29]	QE[2]	"X"[30]	QE[2]	#teeth		QE[2, 27]	QE[2]	
Overjet >9mm or reverse overjet >3.5mm	QE[4, 5]		QE[3, 4, 5]	QE	QE[3, 4, 5]	QE[3, 4, 5]		QE		QE[4, 5]	QE[20]		QE[3, 4, 5]		QE[4, 5]	"X"[4, 5]	QE[3]			QE[4, 5]		QE[20]
Ant. impaction (# teeth)		x3			"X"[1, 37]			QE[24]			QE[3]	QE	QE[24]	"X"[1, 37]	QE[24]	x3[22, 23, 24]	QE[24]	x5			QE[24]	QE[24]
Post impaction (teeth)					"X"[1, 37]			x3[24]				5pts						x5				QE[24]
Ant. Mx crowding >8mm								QE														
Skeletal Class III																				QE		
Congenitally missing teeth																						QE[26]
Crowding/spacing ≥10mm																						QE
Lateral or ant open bite																						QE
Ant/post crossbite of ≥3 teeth/arch																						QE
Point threshold	**26**	**28**	**26**	**15**	**26**	**26**	**28**	**22**	**15**	**26**	**28**	**30**	**26**	**30**	**26**	**30**	**26**	**35**	**26**	**25**	**30**	

QE = qualifying exception; Red = qualifying exception; Black = point-contributing criteria; Green = point-contributing criteria which diverge from the norm

APPENDIX 1b: Point-contributing conditions utilization among state-modified HLD indices, with weight factors

	AK	*AR*	*CA*	*DC*	*DE*	*FL*	*IL*	*MA*	*MD*	*ME*	*MO*	*MT*	*NJ*	*NM*	*NY*	*OK*	*RI*	*SC*	*TX*	*WA*	*WY*
Anterior Open bite (mm)	x4	x4	x4[3]	X4	x4[3]	x4[3]	x4[3]	x4	x4	x4	x4[3]	x4	x4[3]	x4	x4	x4	x4	x4	x4	x4[3]	x4
Overjet in mm	x1	**x2**	x1[3]	x1[18]	x1[3]	x1[3, 18]	x1[3, 36]	x1	x1[18]	x1	x1[3]	x1	x1[3]	x1	x1	x1[19]	x1[3]	**x3**	x1[18]	x1[3]	x1
Overbite in mm	x1[8]	x1[8]	x1[8]	x1[17]	x1[3, 8]	x1[3, 17]	x1[3, 8]	x1[8]	x1[17]	x1[8]	x1[8]	x1[8]	x1[8]	x1[8]	x1[8]	x1[8]	x1[8]	**x2**	x1[8, 17]	x1[3, 8]	x1[8]
Md protrusion (mm)	x5	x5[9]	x5[3]	x5[10]	x5[3]	x5	x5[3, 36]	x5[10]	x5	x5	x5	x5	x5[3]	x5	x5	x5	x5	x5	x5	x5[3]	x5
Ant ectopia (teeth)	x3[6]	x3[6, 12]	x3[6, 13]	x3[6]	x3[6, 13]	x3[6]	x3[6]	x3[6]	x3[6]	x3[6]	x3[6]	x3[6]	x3[6, 13]	x3[6]	x3[6, 13]	x3[6]	x3[6]		x3[6]	x3[6, 13]	x3[6]
Ant crowding (arches)	x5[6, 7]	x5[6,15]	x5[6, 7]	x5[6, 7]	x5[6, 7]	x5[6, 7]	x5[6, 7]	x5[6, 7]	x5[6, 7, 14]	x5[6, 7]	x5[6, 7]	x5[6, 7]	x5[6, 7]	x5[6, 7]	x5[6, 7]	x5[6, 7]	x5[6, 7]		x5[6, 7]	x5[6, 7]	x5[15]
Post unilateral crossbite	4pts 33, 35		4pts 33, 35	**5pts**	4pts 33, 35	4pts 33, 35	4pts 33, 34	4pts 33, 35	**5pts**	4pts 33, 35	4pts 33, 35	4pts 33, 35	4pts 33, 35	4pts 33, 35	4pts 33, 35	4pts 33, 35	4pts 33, 35	**#teeth[34]**		4pts 33, 35	4pts 33, 35
Post ectopia (teeth)	x3[6]		x3[6, 13]	x3[6]	x3[6, 13]	x3[6]	x3[6]	x3[6]	x3[6]	x3[6]	x3[6]	x3[6]	x3[6, 13]	x3[6]	x3[6, 13]	x3[6]	x3[6]			x3[6, 13]	x3
Labiolingual spread	x1	x1[32]	x1	x1[32]	x1	x1	x1	x1[16]	x1[32]	x1	x1	x1	x1	x1	x1				x1[32]		
Post bilateral crossbite					4pts 33, 35		4pts 33, 34		**5pts**		4pts 33, 34	**8pts 33, 35**				4pts 33, 35		**#teeth[34]**		**8pts 33, 35**	
Ant. impaction (# teeth)		**x3**			"X"[1, 37]			QE[24]			QE[3]	QE	QE[24]	"X"[1, 37]	QE[24]	**x3[22, 23, 24]**	QE[24]	**x5**			QE[24]
Post impaction (teeth)					"X"[1, 37]			**x3[24]**				**5pts**						**x5**			
Congenitally missing posterior teeth								x3[25]													
Crowding <6mm (arches)																		x2			
Crowding >6mm (arches)																		x4			
Ant spacing (mm)																x1					
Midlines (mm)																x1					
Functional shift of mandible (as a result of posterior crossbite)																4pts					
Habits affecting arch development																		#habits x2			
Psychological factors affecting development													"X"								
Point threshold	**26**	**28**	**26**	**15**	**26**	**26**	**28**	**22**	**15**	**26**	**28**	**30**	**26**	**30**	**26**	**30**	**26**	**35**	**26**	**25**	**30**

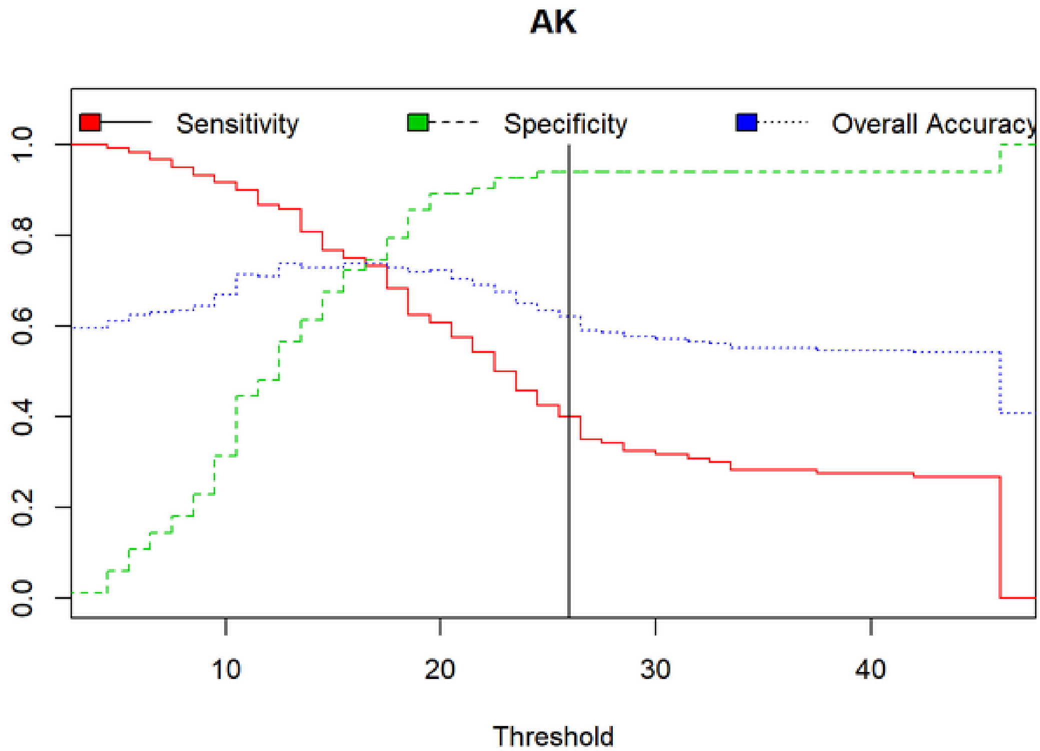
QE = qualifying exception; Red = qualifying exception; Black = point-contributing criteria; Green = point-contributing criteria which diverge from the norm

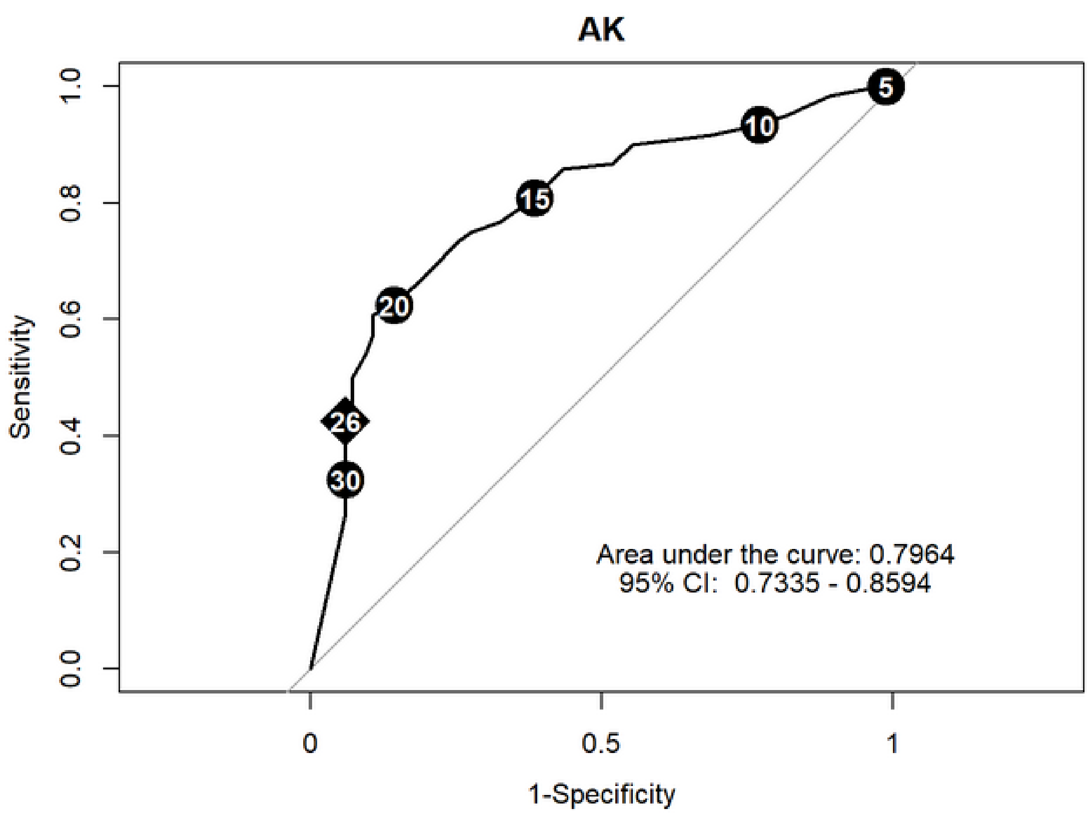
Legend for Appendix 1a-b

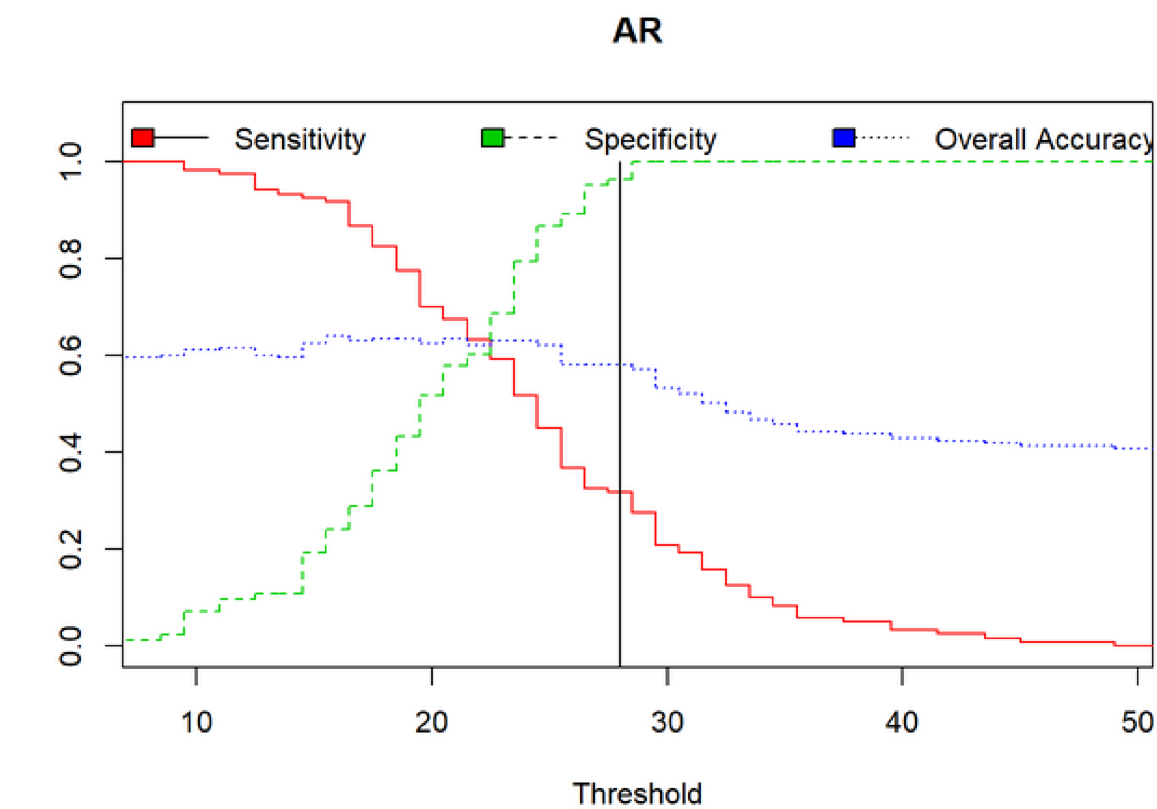
1 Not on HLD index, but in provider manual; does not include search of state legislature
2 Tissue laceration and/or clinical attachment loss must be present
3 Measured at the maxillary central incisors ONLY
4 With incompetent lips
5 With reported masticatory/speech difficulties
6 Anterior crowding OR ectopic eruption - most severe condition only
7 Anterior crowding greater than 3.5mm
8 Count reverse overbite
9 Class III case only
10 Measured from the molar MB groove to the MB cusp of the maxillary molar
11 2mm or more of 4 or more teeth per arch in anterior OR lateral open bite
12 Impacted teeth included
13 Must be greater than 50% blocked out
14 Anteriors so crowded that extractions are prerequisite
15 Anterior crowding greater than 0mm
16 Including anterior spacing
17 Subtract 3mm
18 Subtract 2mm
19 2+ teeth must be involved
20 Equal to or greater
21 Or multiple mandibular teeth touching the palate
22 Contralateral tooth has erupted 12months ago
23 Impacted tooth apex must be closed
24 Treatment planned to be brought into occlusion
25 Posterior teeth only
26 2 or more teeth of at least one tooth per quadrant
27 Clinical attachment loss must be greater than 1mm
28 Clinical attachment loss must be greater than 1.5mm
29 Causing impingement
30 Must present with tooth damage, anterior tooth which is completely lingual, or destruction of soft tissue or gingival recession
31 Anterior and/or posterior crowding; Categories included moderate crowding (<6mm), severe crowding (>6mm)
32 Spacing in mm of anterior teeth only
33 At least one molar
34 1 or more teeth must be involved
35 2 or more adjacent teeth must be involved
36 Only count overjet or mandibular protrusion, whichever is greatest
37 Causes damage to other permanent tooth root
38 Includes cleft lip

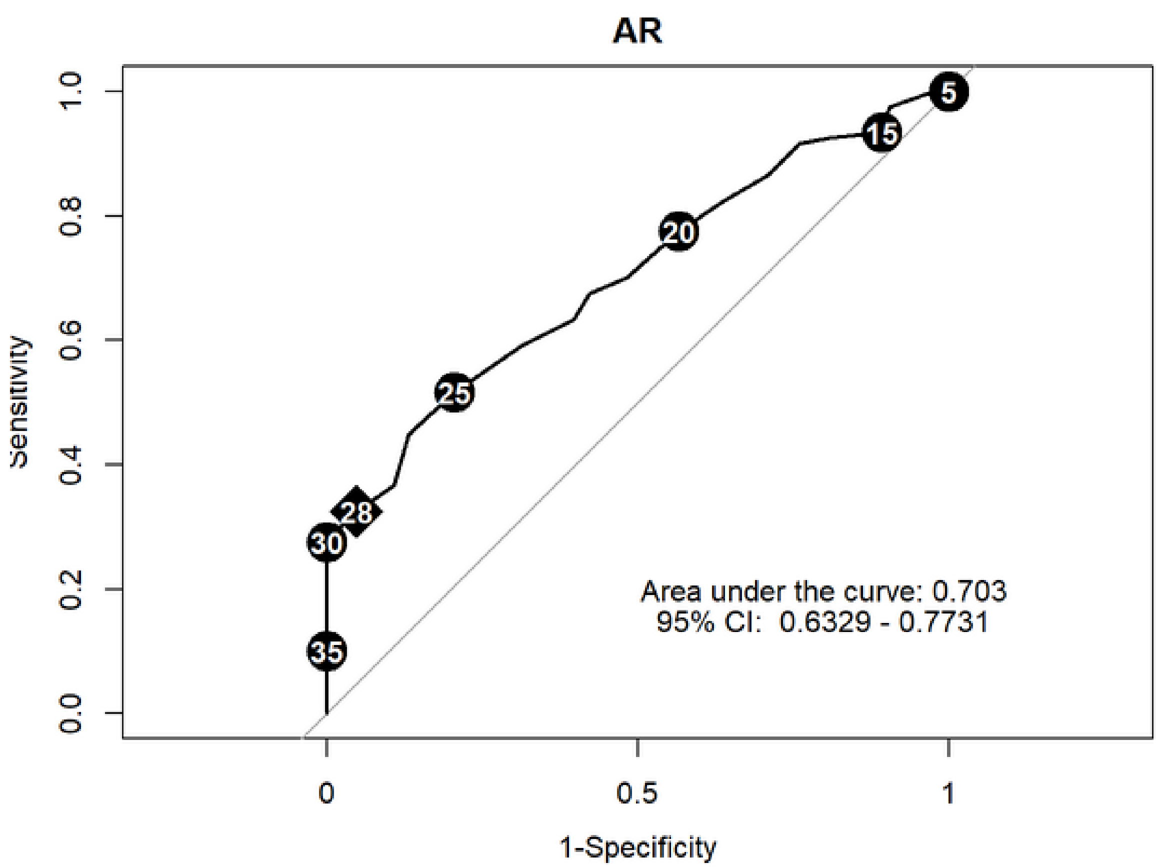
APPENDIX 2: ROC curves and sensitivity/specificity/accuracy plots for HLD-using States

Vertical line in sensitivity/specificity/accuracy plots and diamond marker in ROC curves indicate current point threshold.

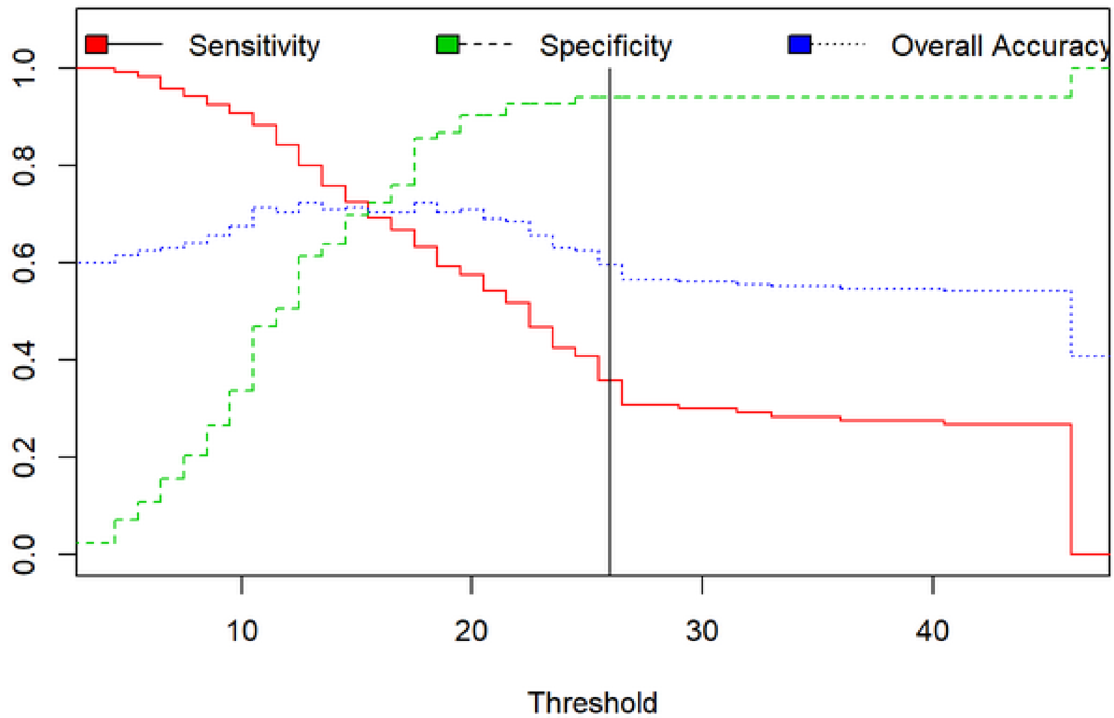




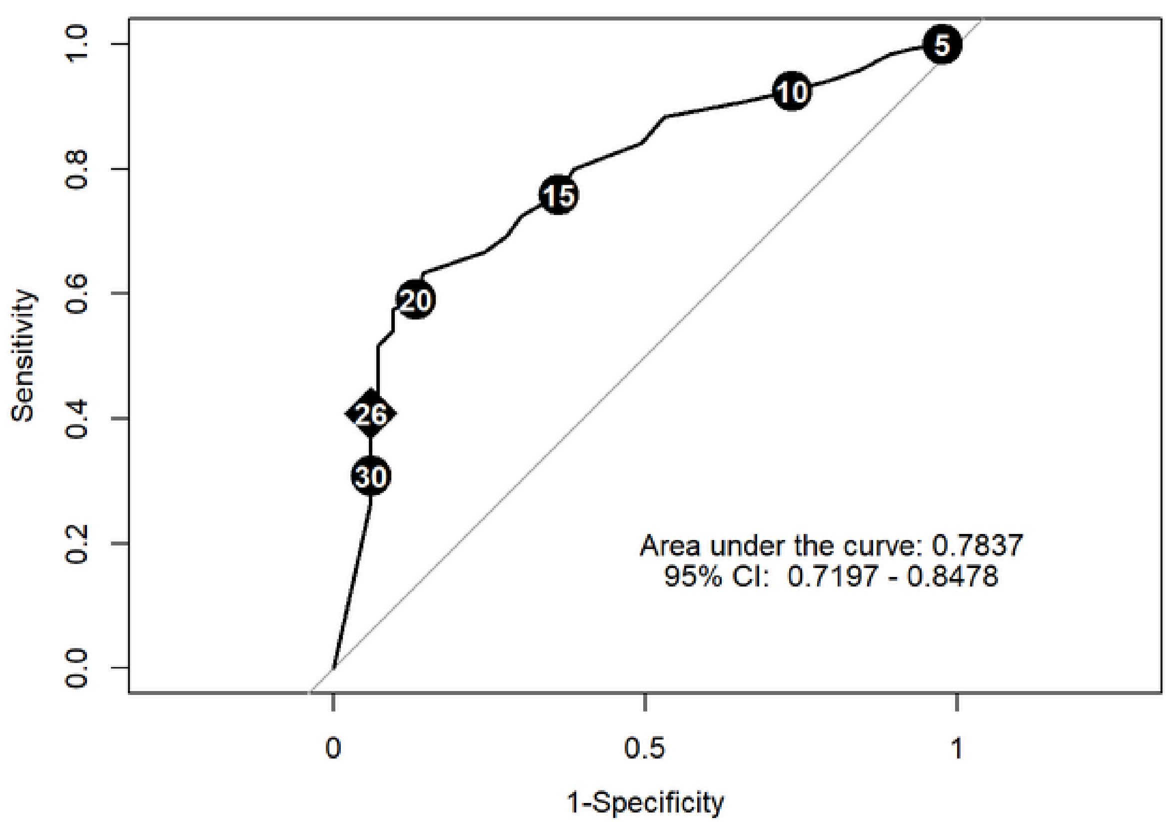

AR
Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
40
50
Threshold

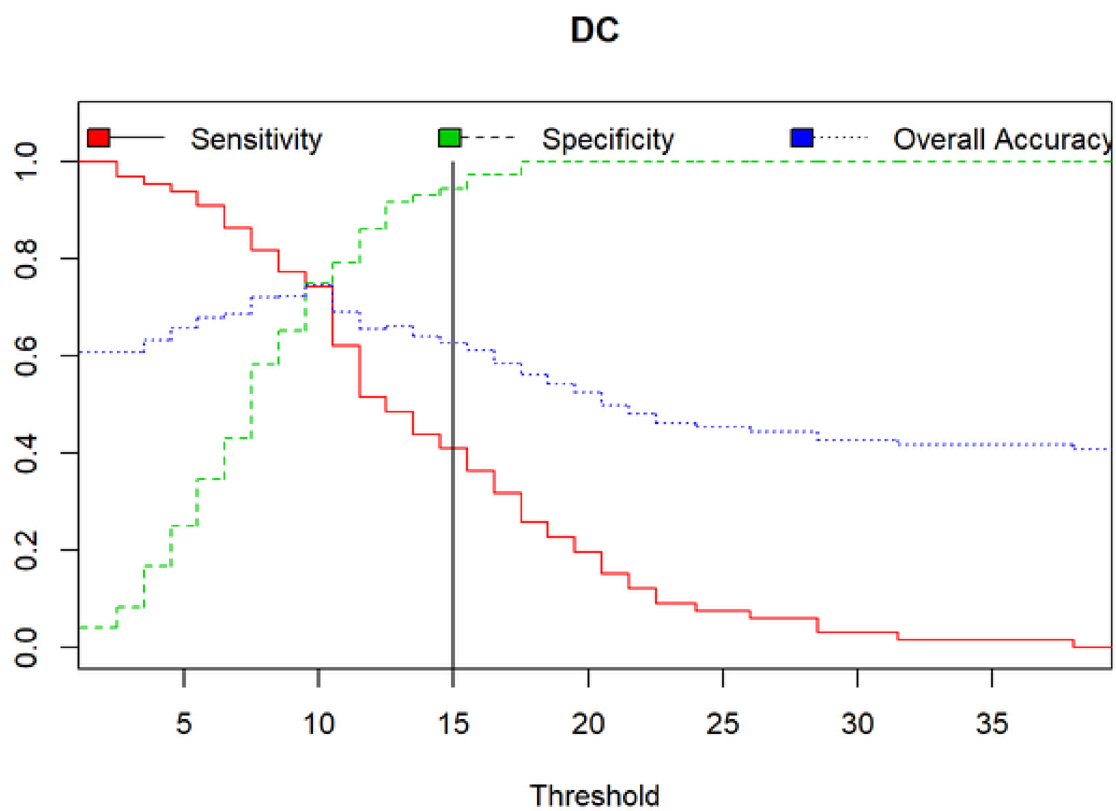

AR
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
5
15
20
25
28
30
35
Area under the curve: 0.703
95% CI: 0.6329 - 0.7731
0
0.5
1
1-Specificity

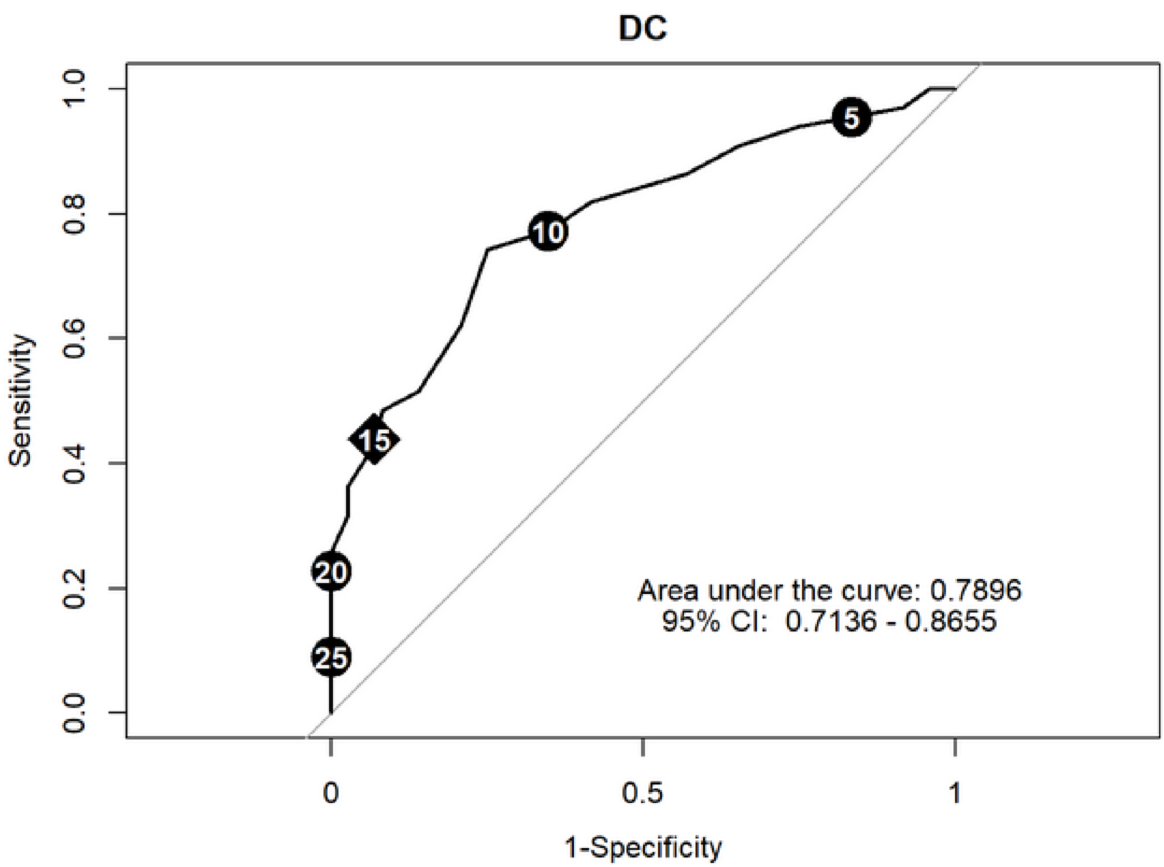
CA

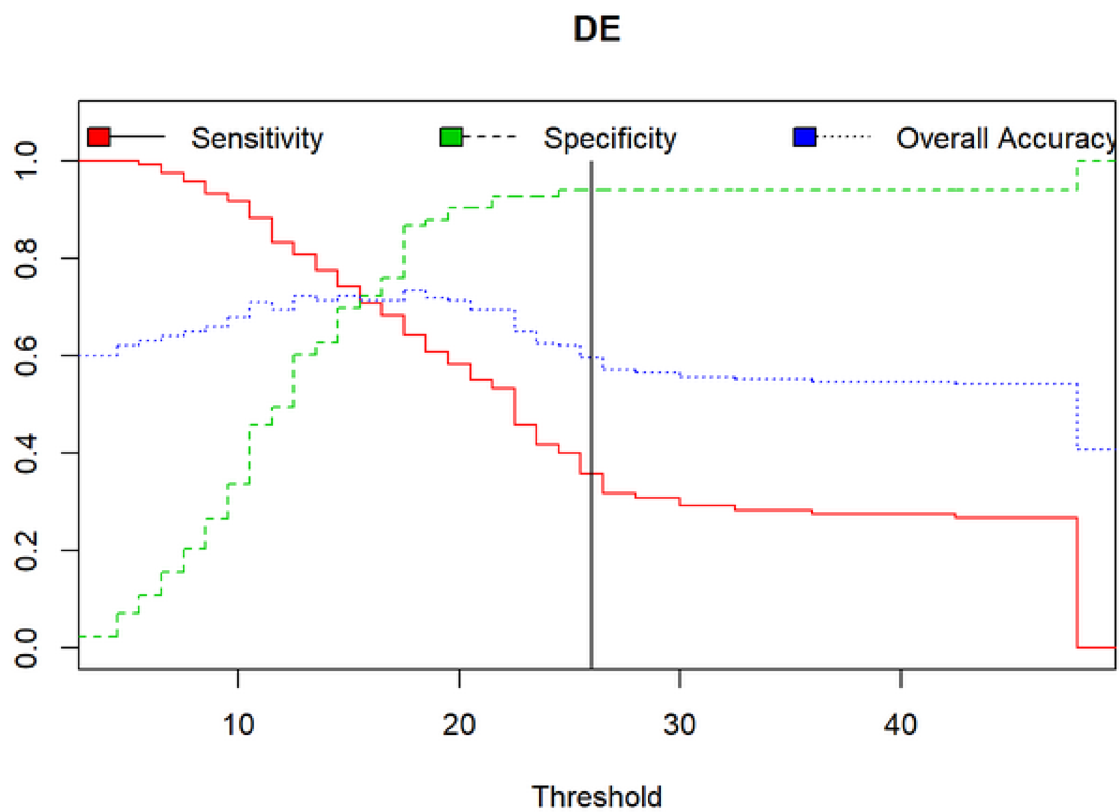

Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
40
Threshold

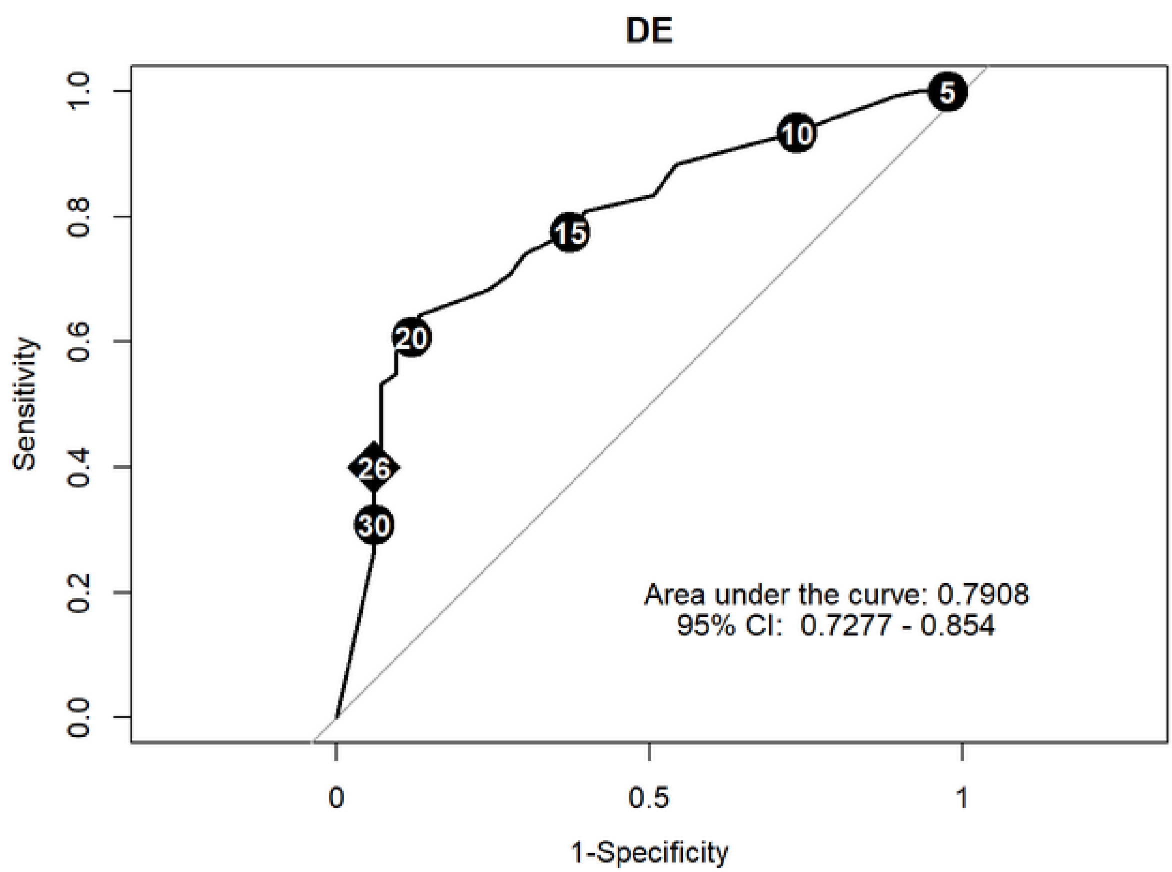
CA


Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
26
30
Area under the curve: 0.7837
95% CI: 0.7197 - 0.8478
0
0.5
1
1-Specificity

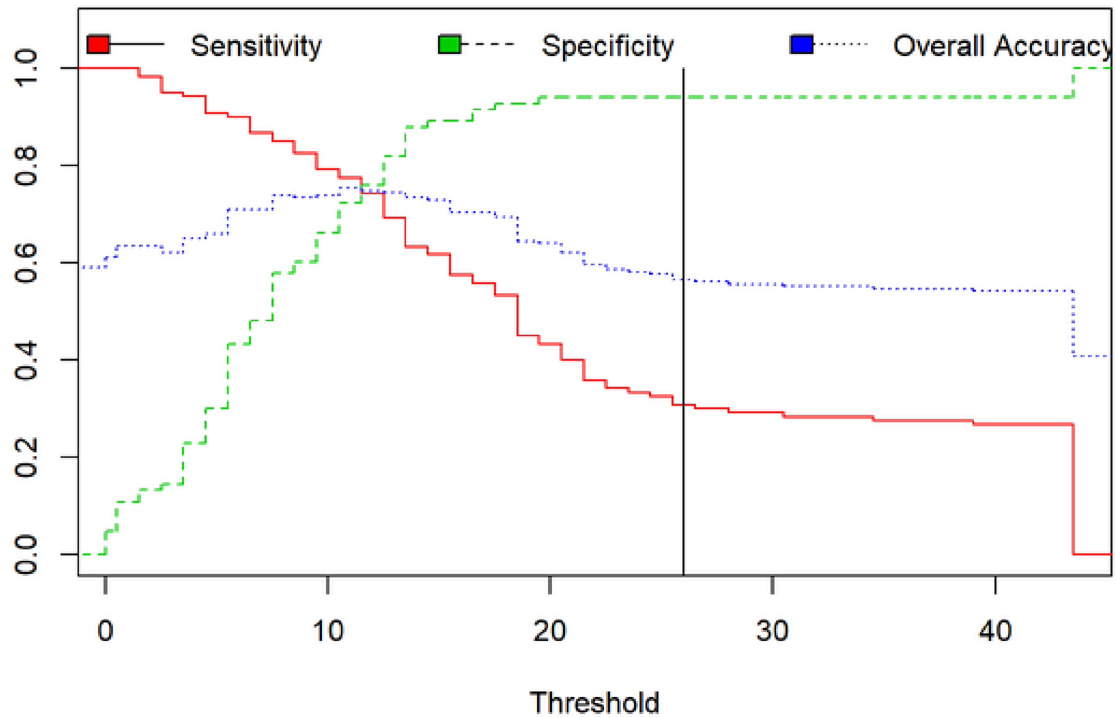

DC
Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
25
30
35
Threshold


DC
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
25
Area under the curve: 0.7896
95% CI: 0.7136 - 0.8655
0
0.5
1
1-Specificity

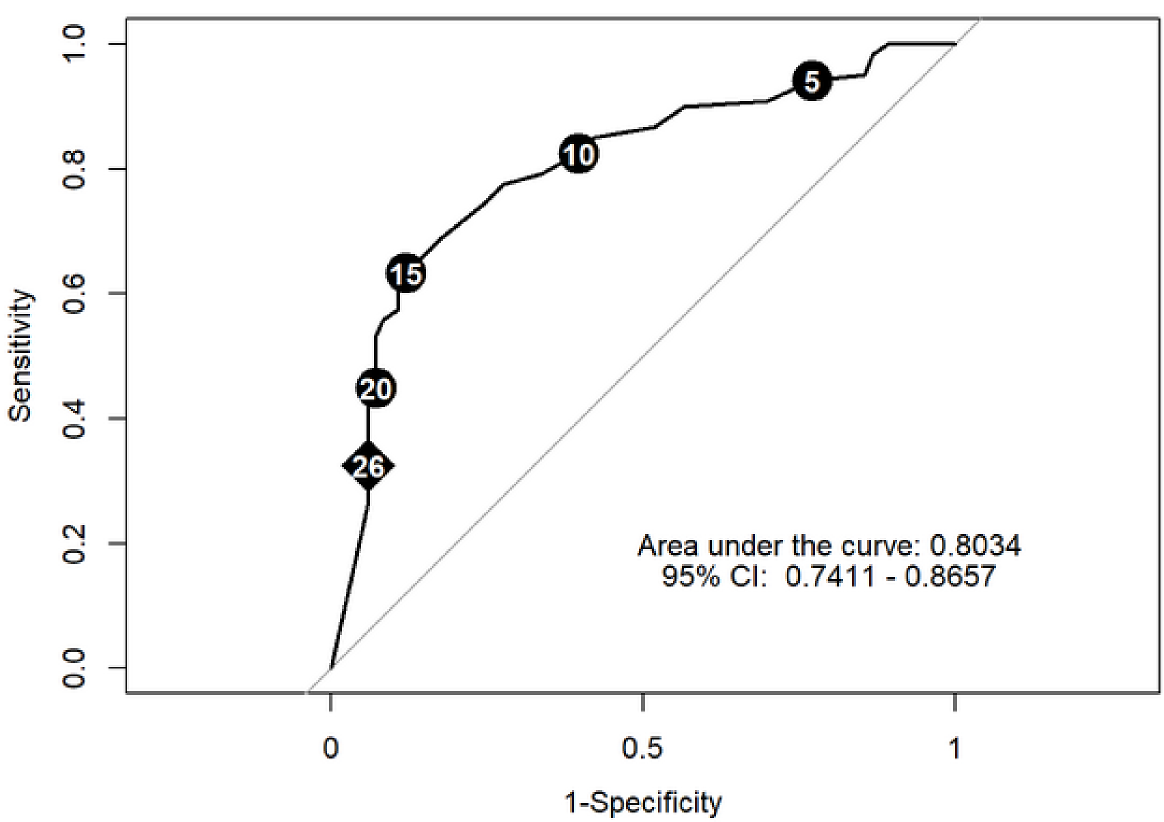

DE
Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
40
Threshold

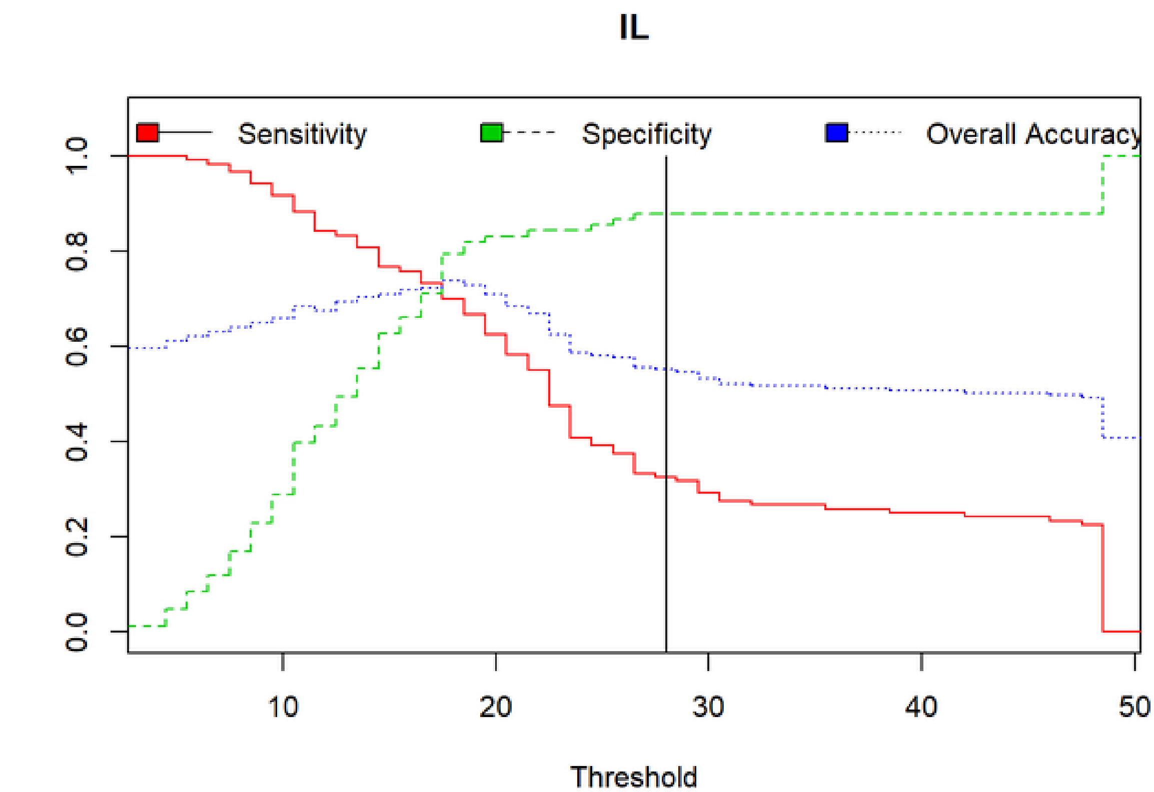

DE
Sensitivity
1-Specificity
1.0
0.8
0.6
0.4
0.2
0.0
0
0.5
1
5
10
15
20
26
30
Area under the curve: 0.7908
95% CI: 0.7277 - 0.854

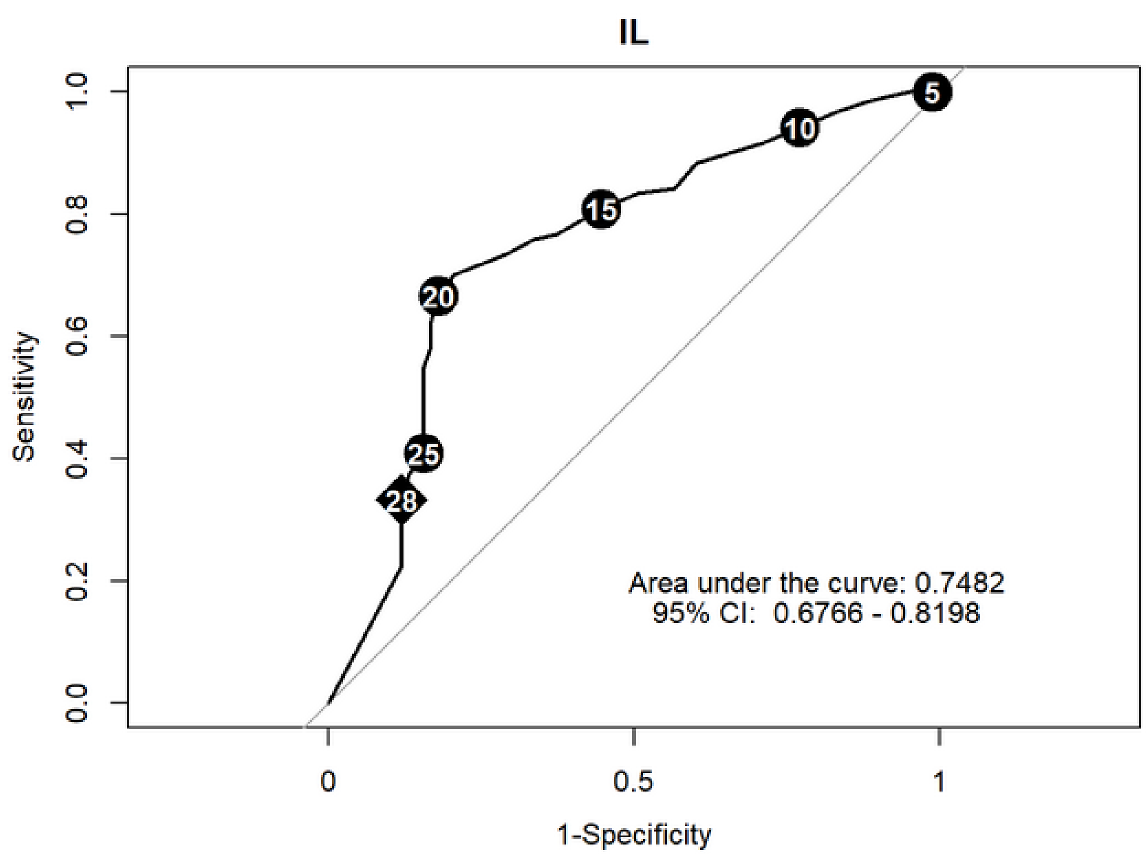
FL

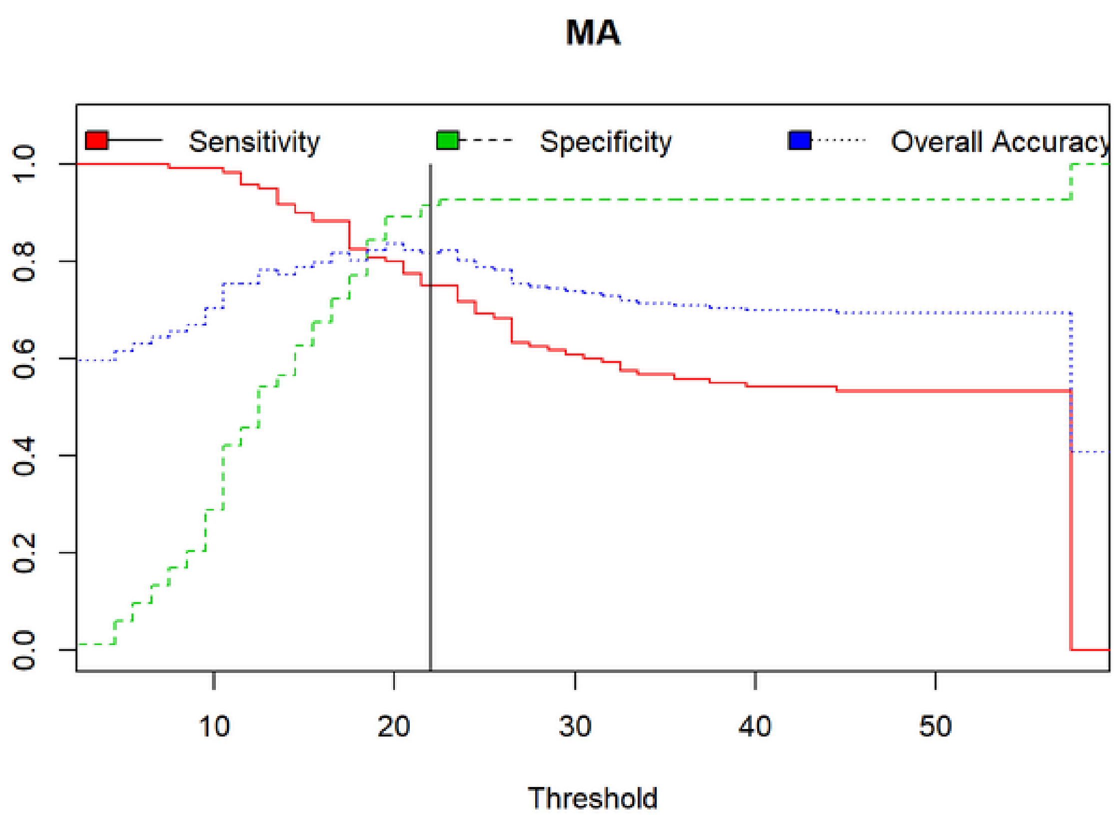

Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
Threshold

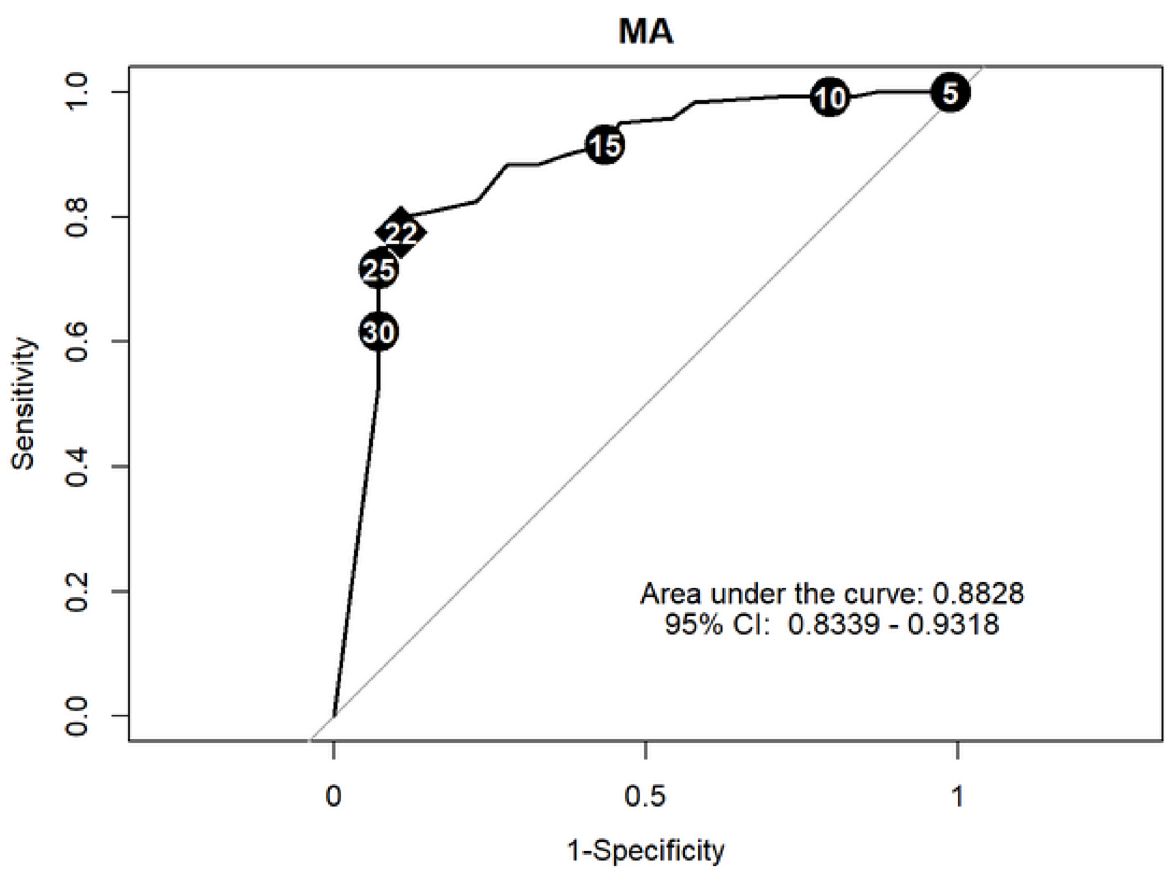
FL

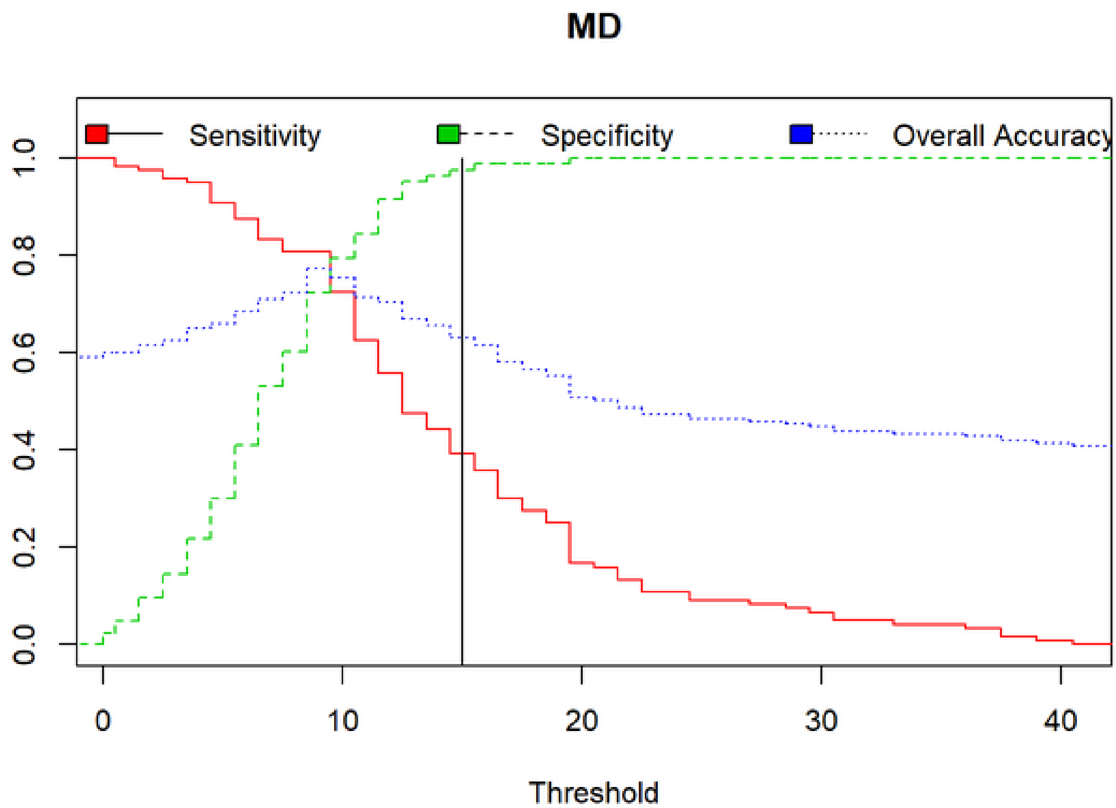

Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
26
Area under the curve: 0.8034
95% CI: 0.7411 - 0.8657
0
0.5
1
1-Specificity

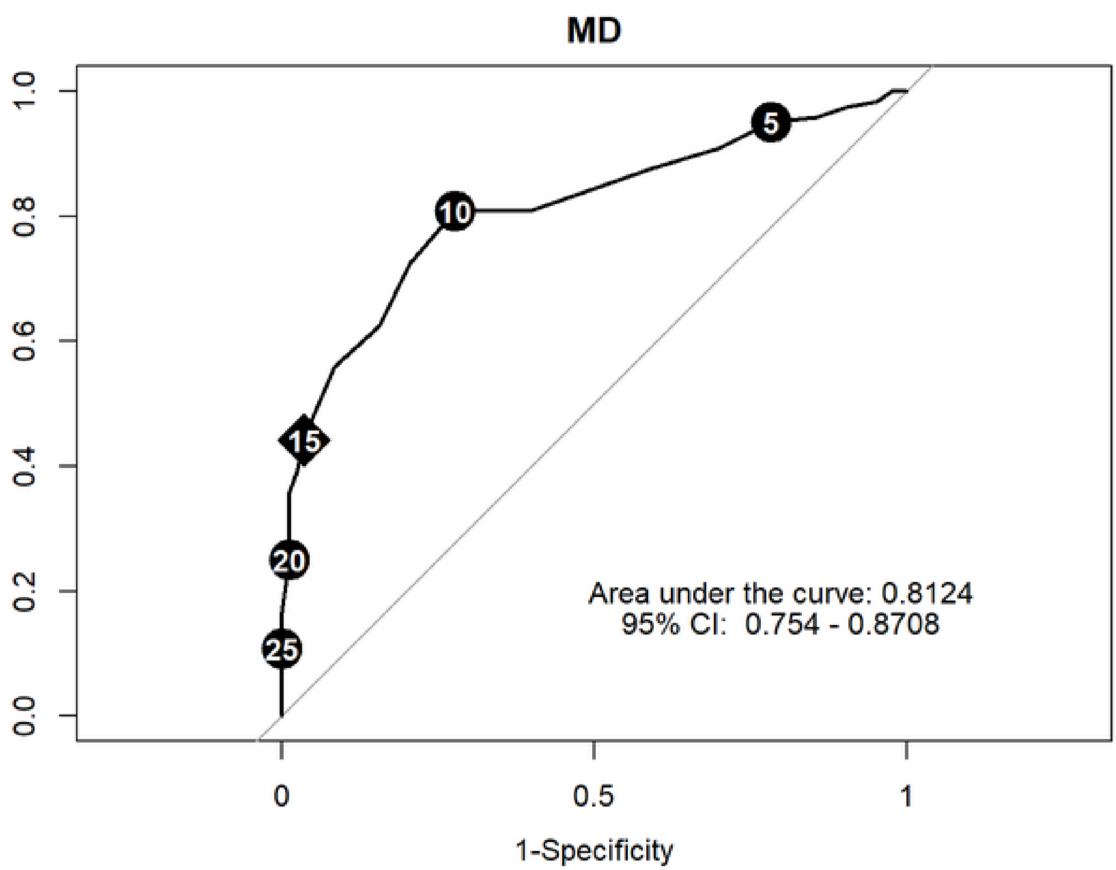

IL
Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
40
50
Threshold

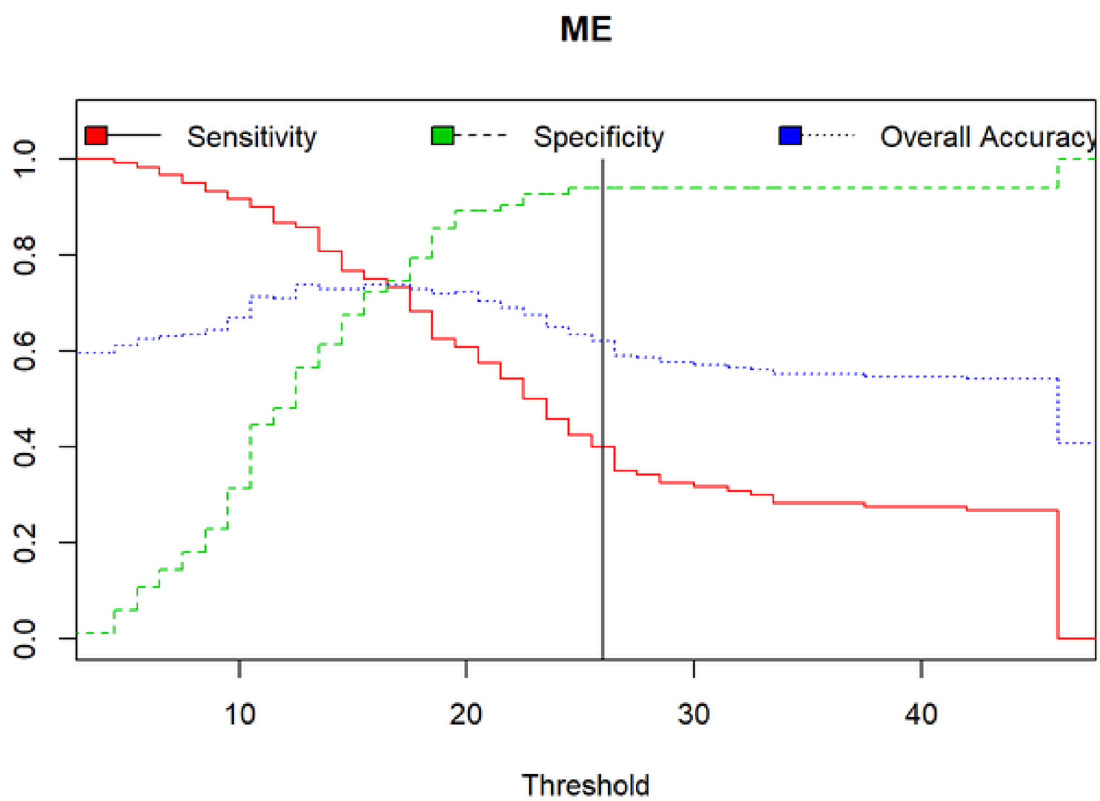

IL
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
25
28
Area under the curve: 0.7482
95% CI: 0.6766 - 0.8198
0
0.5
1
1-Specificity

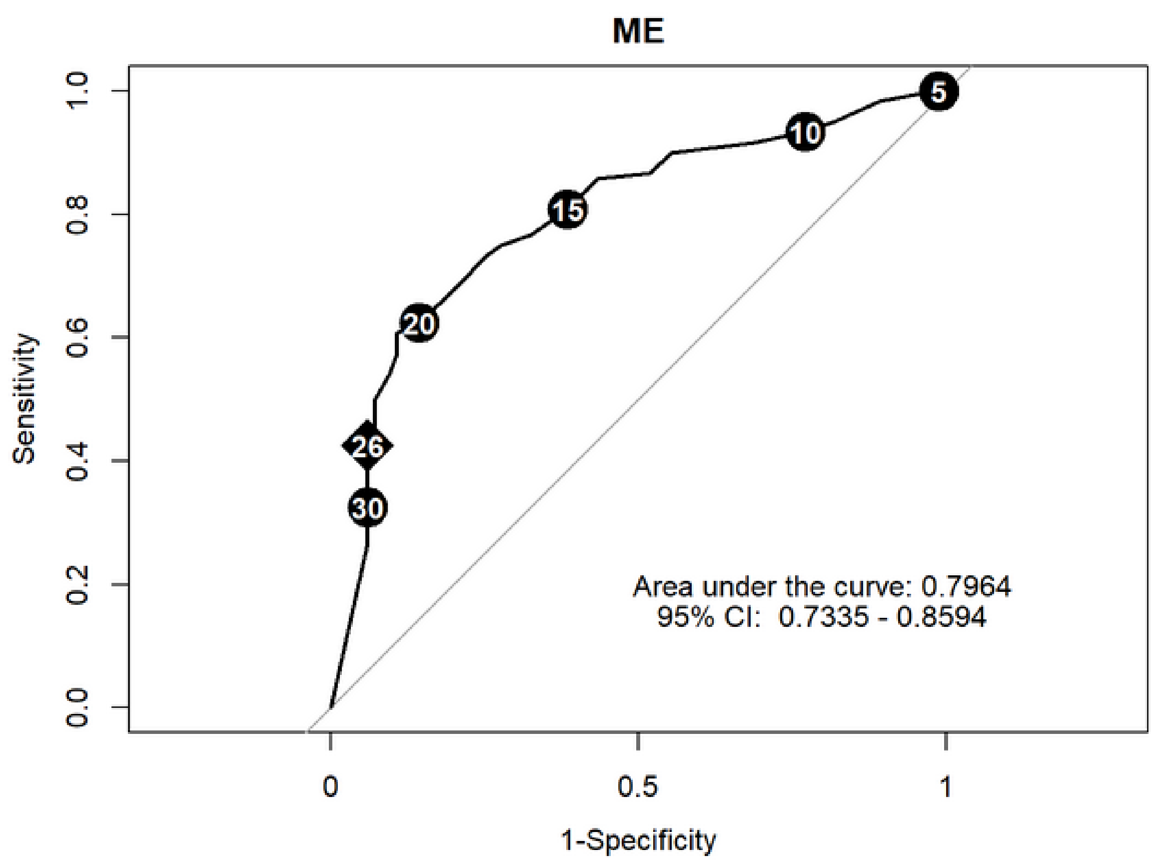

MA
Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
40
50
Threshold

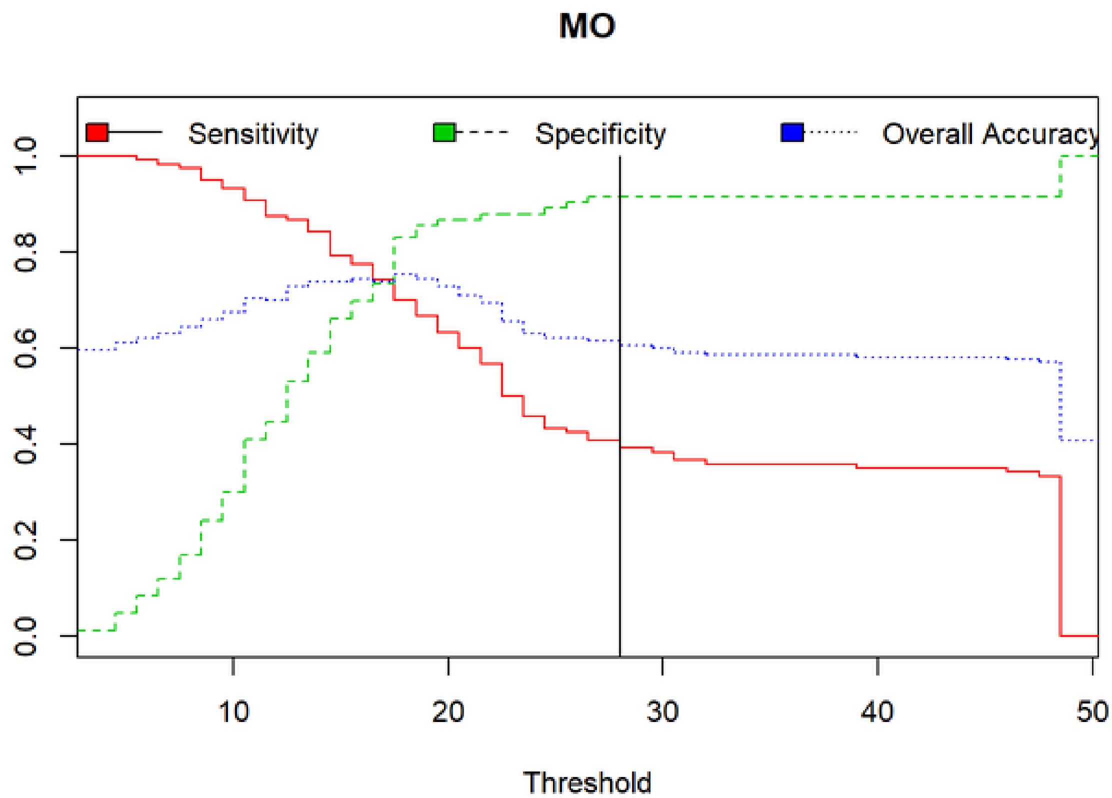

MA
Sensitivity
1-Specificity
1.0
0.8
0.6
0.4
0.2
0.0
0
0.5
1
5
10
15
22
25
30
Area under the curve: 0.8828
95% CI: 0.8339 - 0.9318

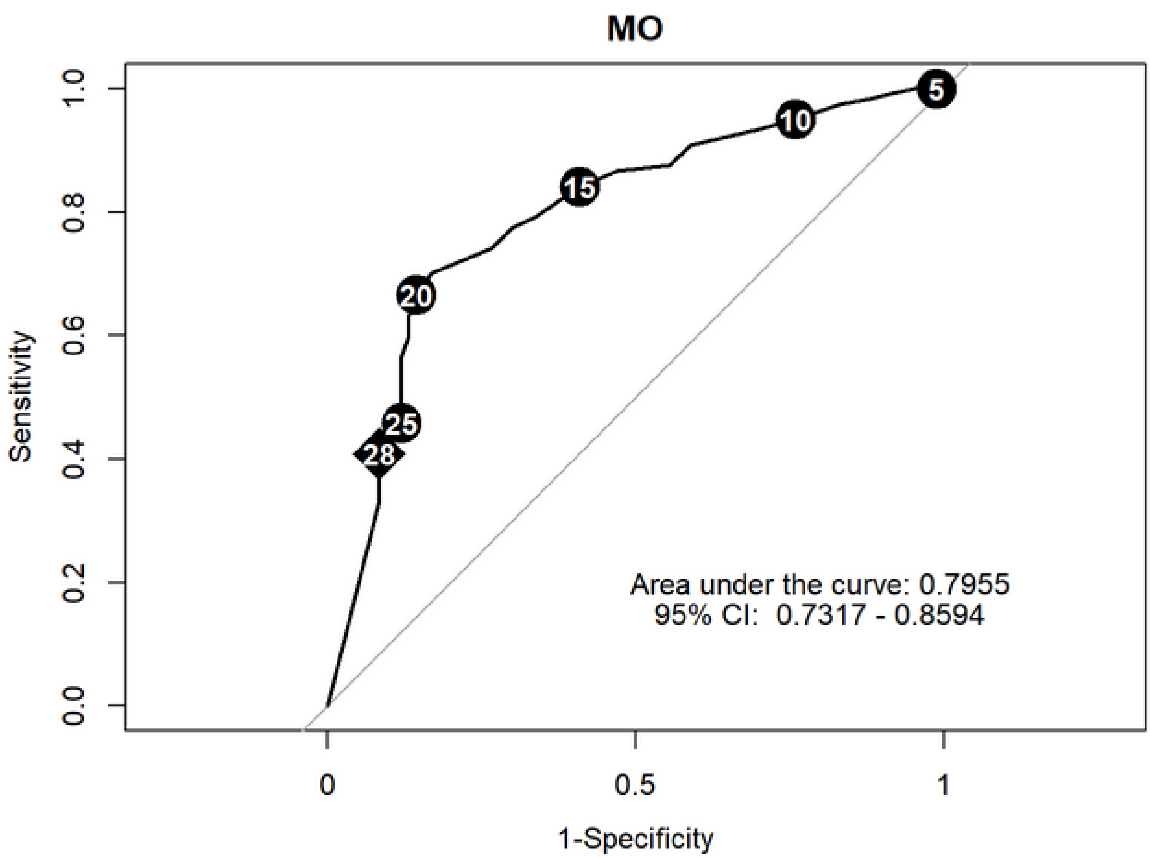

MD
Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
Threshold

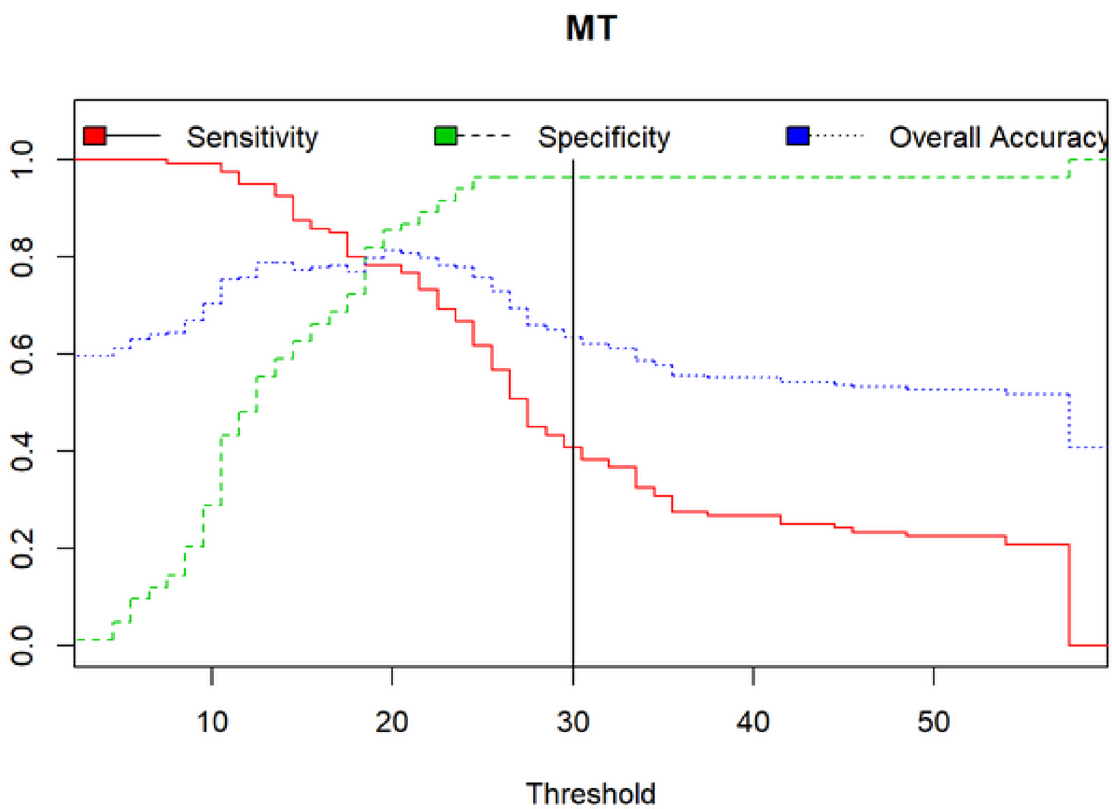

MD
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
25
Area under the curve: 0.8124
95% CI: 0.754 - 0.8708
0
0.5
1
1-Specificity

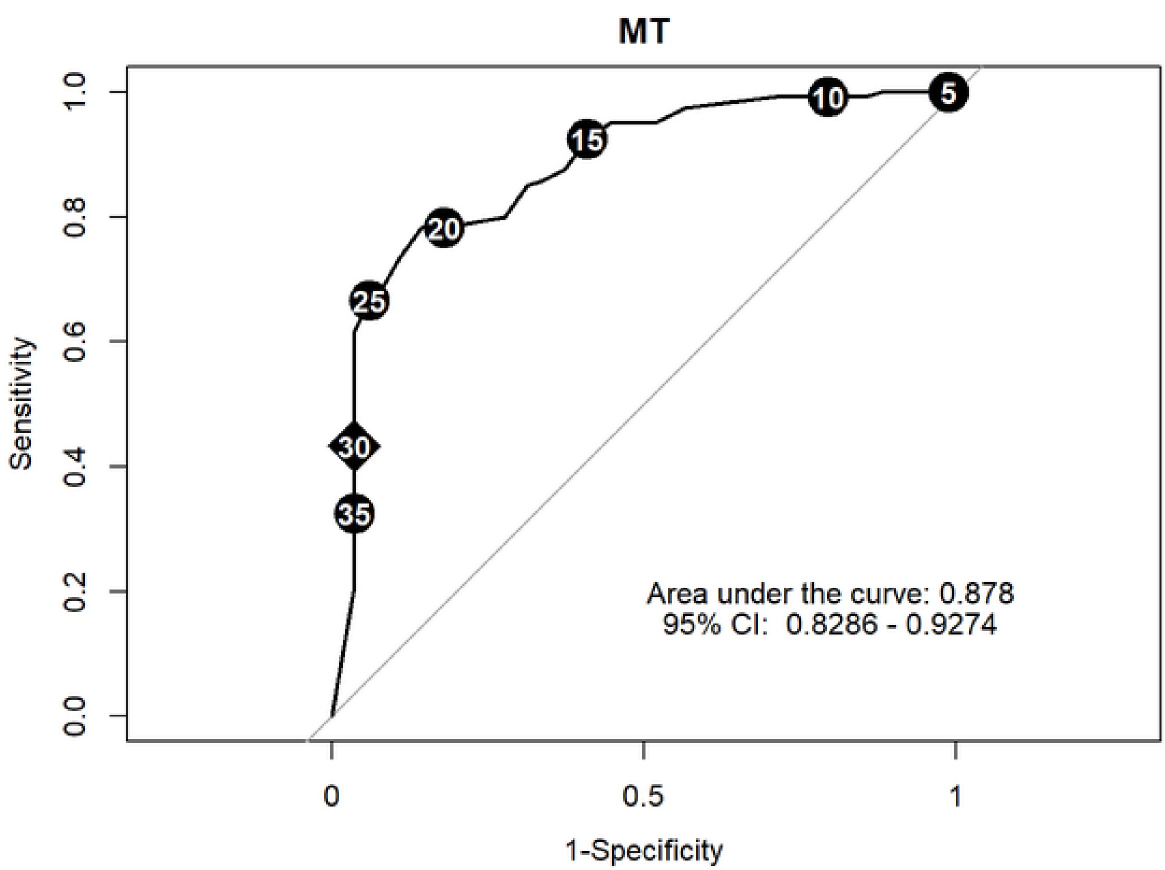

ME
Sensitivity
Specificity
Overall Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
10
20
30
40
Threshold

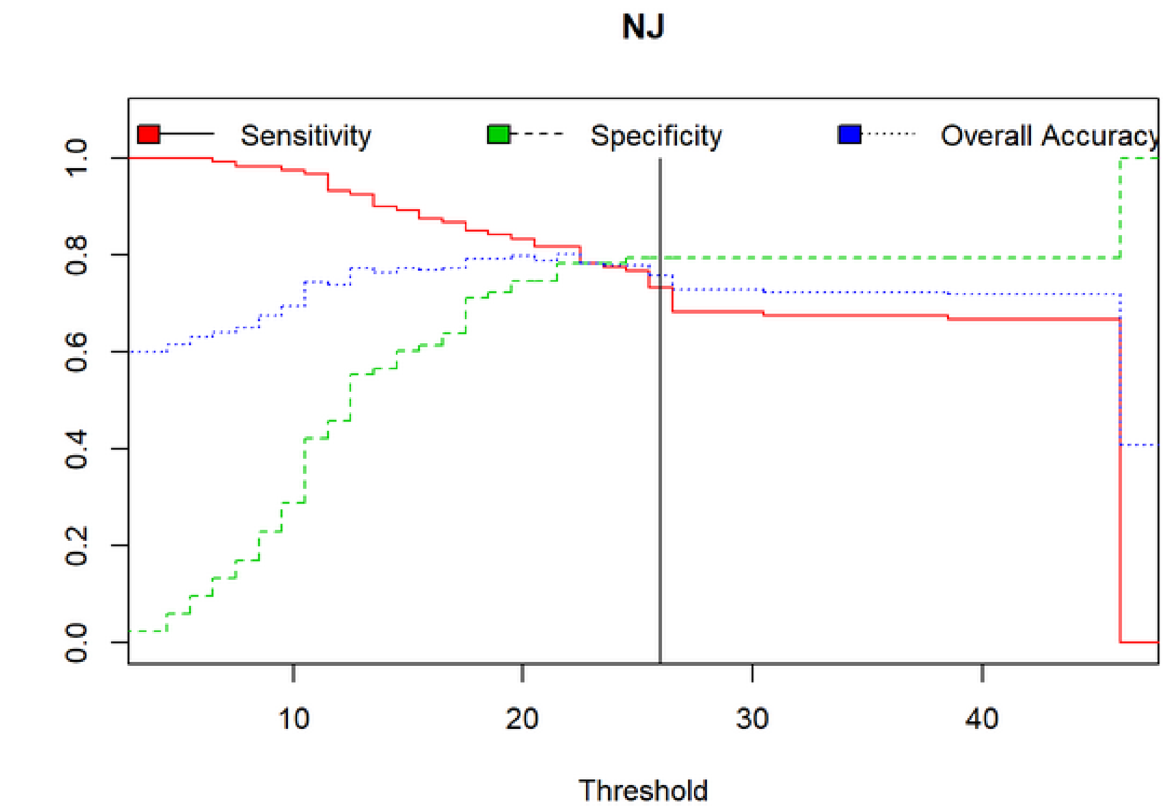

ME
Sensitivity
1-Specificity
0.0
0.2
0.4
0.6
0.8
1.0
0
0.5
1
5
10
15
20
26
30
Area under the curve: 0.7964
95% CI: 0.7335 - 0.8594

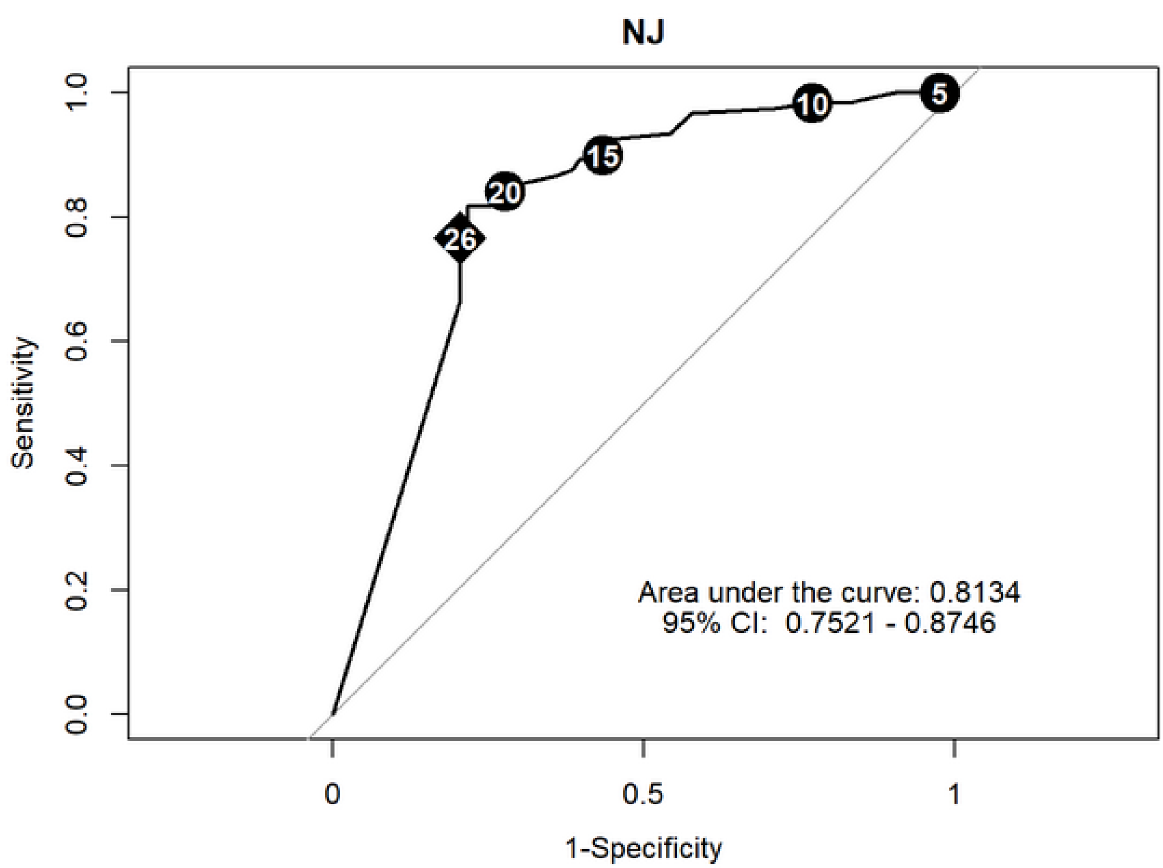

MO
Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
40
50
Threshold

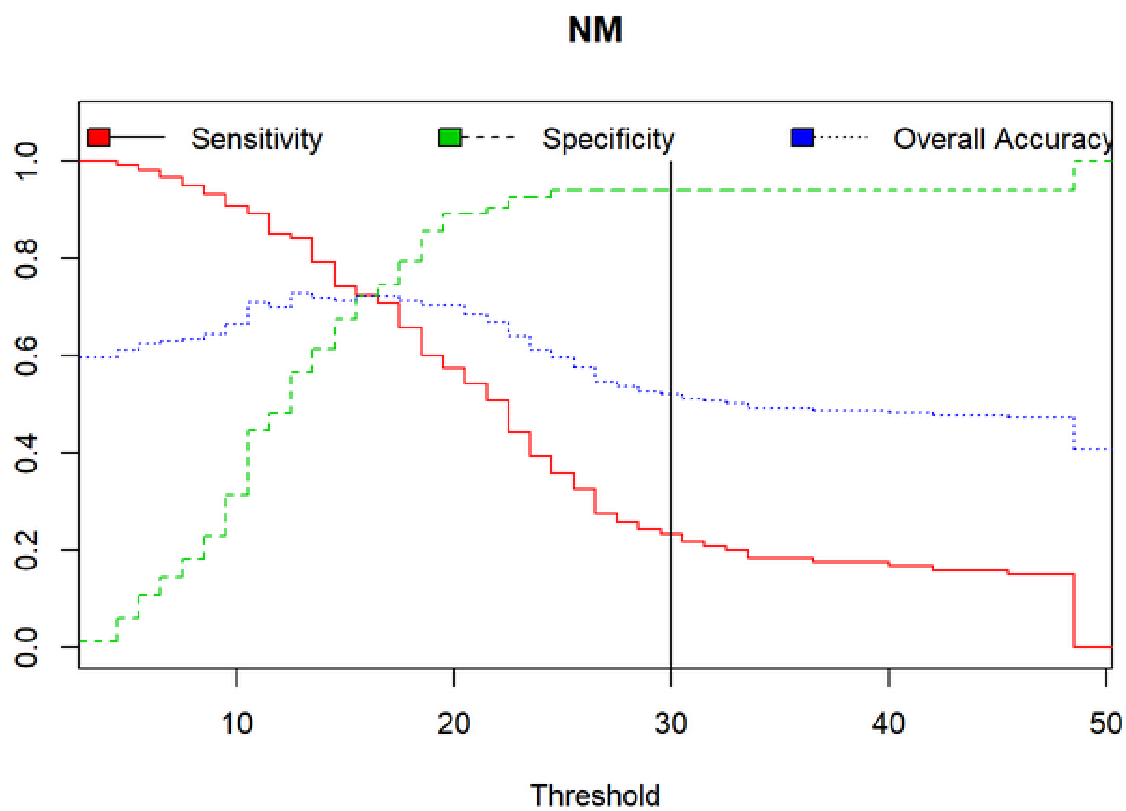

MO
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
25
28
Area under the curve: 0.7955
95% CI: 0.7317 - 0.8594
0
0.5
1
1-Specificity

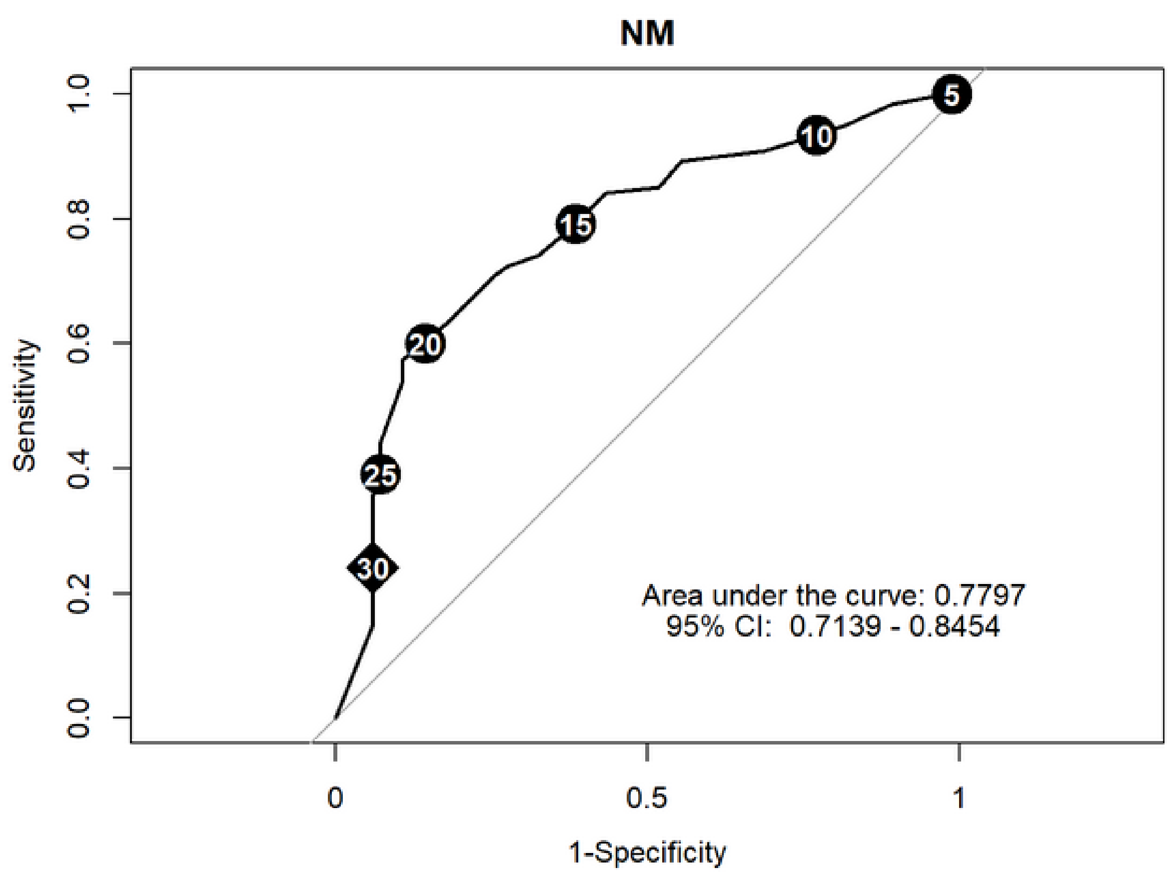

MT
Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
40
50
Threshold

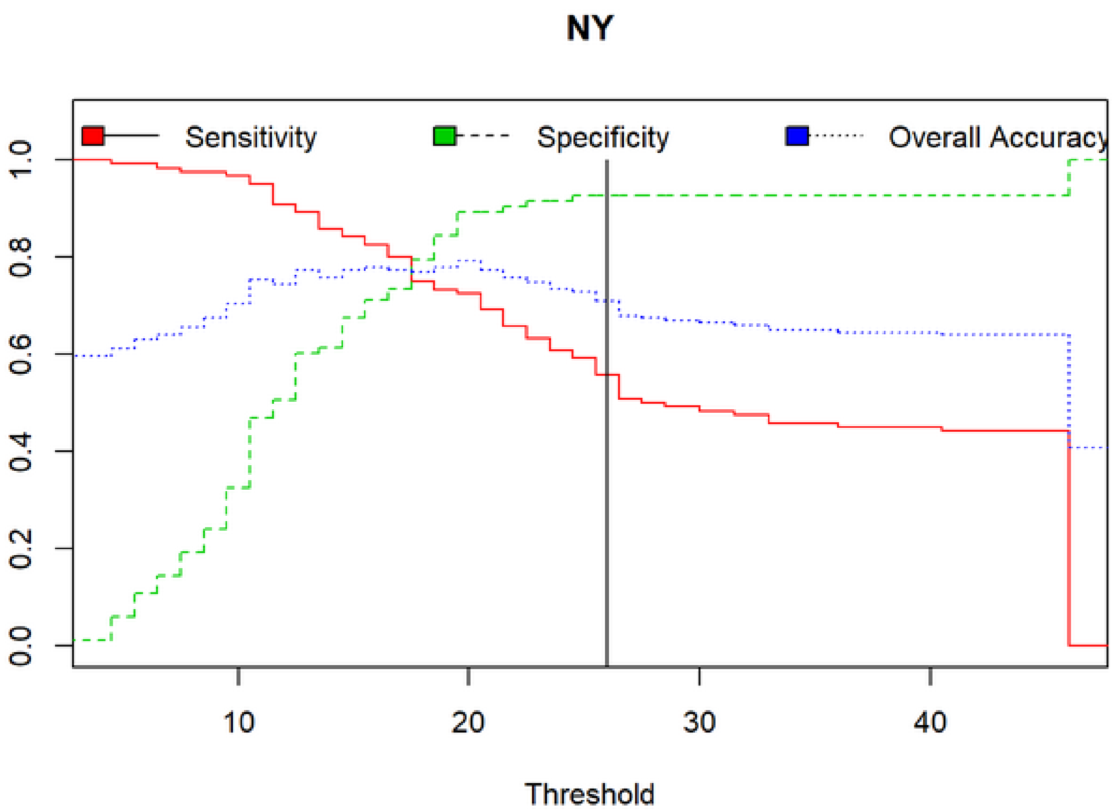

MT
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
25
30
35
Area under the curve: 0.878
95% CI: 0.8286 - 0.9274
0
0.5
1
1-Specificity

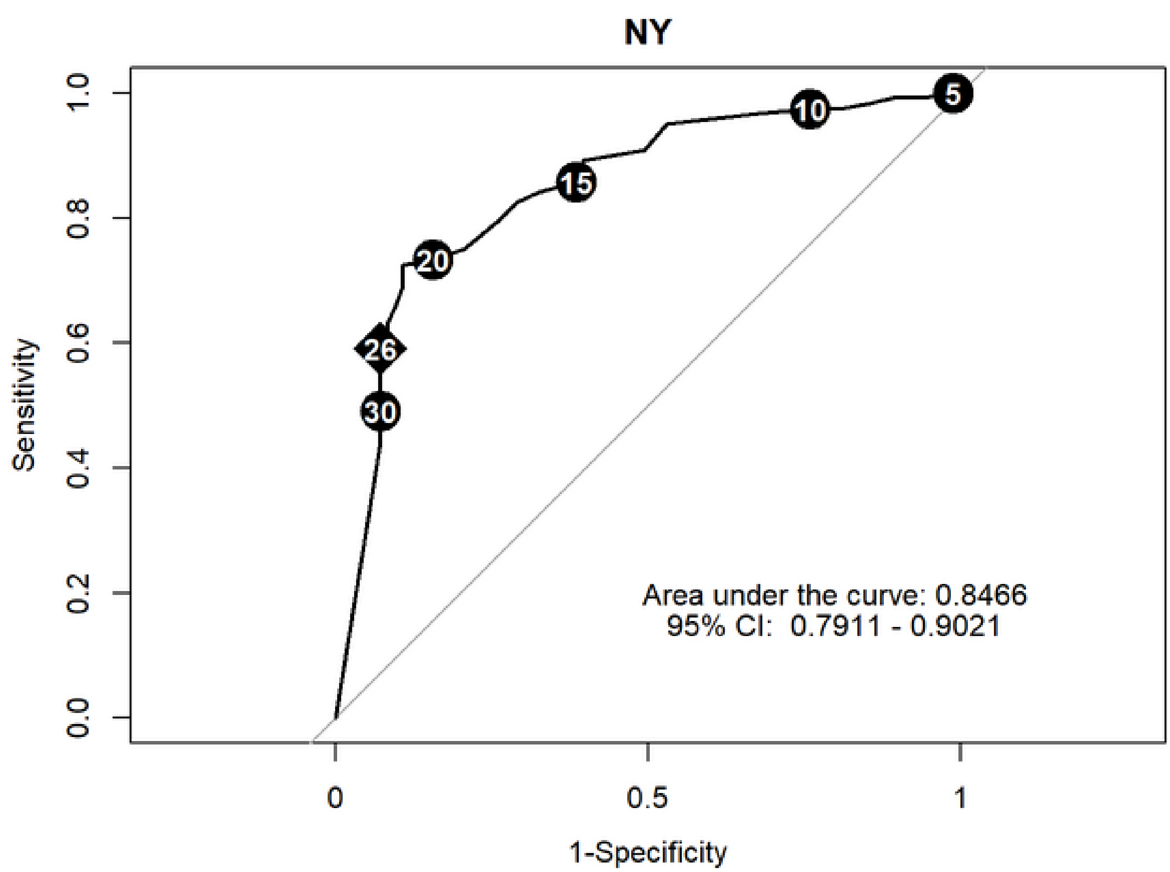

NJ
Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
40
Threshold

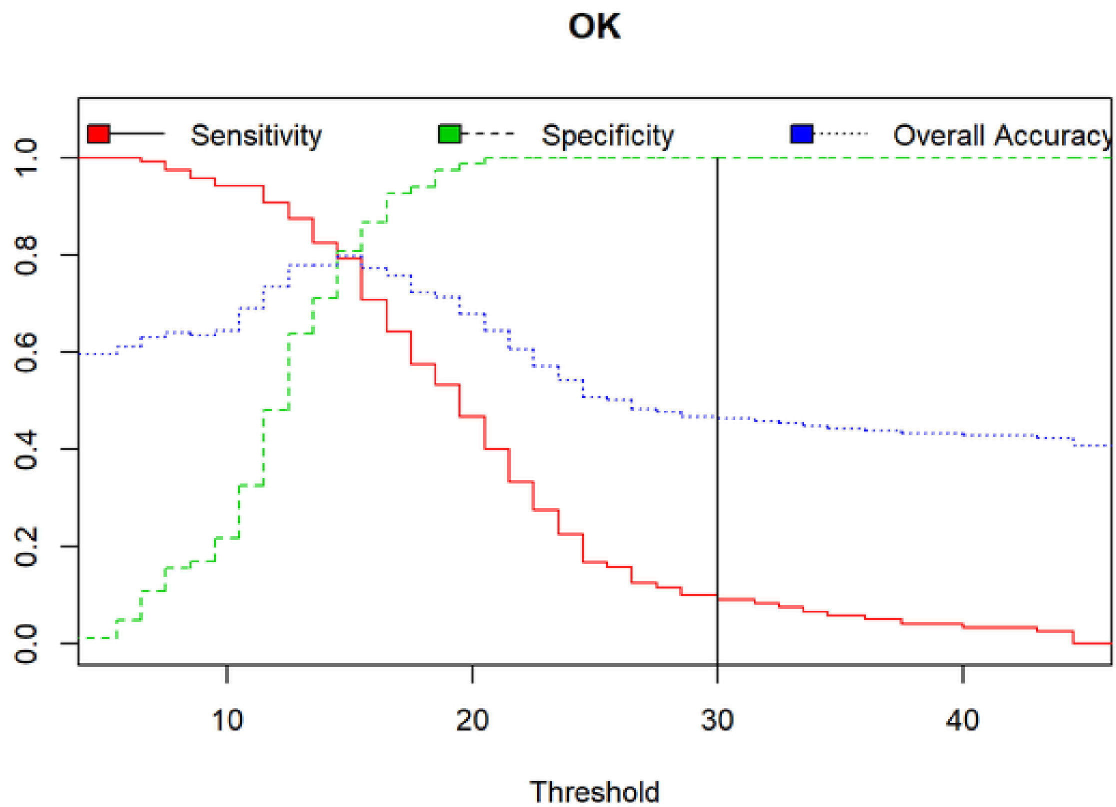

NJ
Sensitivity
1-Specificity
1.0
0.8
0.6
0.4
0.2
0.0
0
0.5
1
5
10
15
20
26
Area under the curve: 0.8134
95% CI: 0.7521 - 0.8746

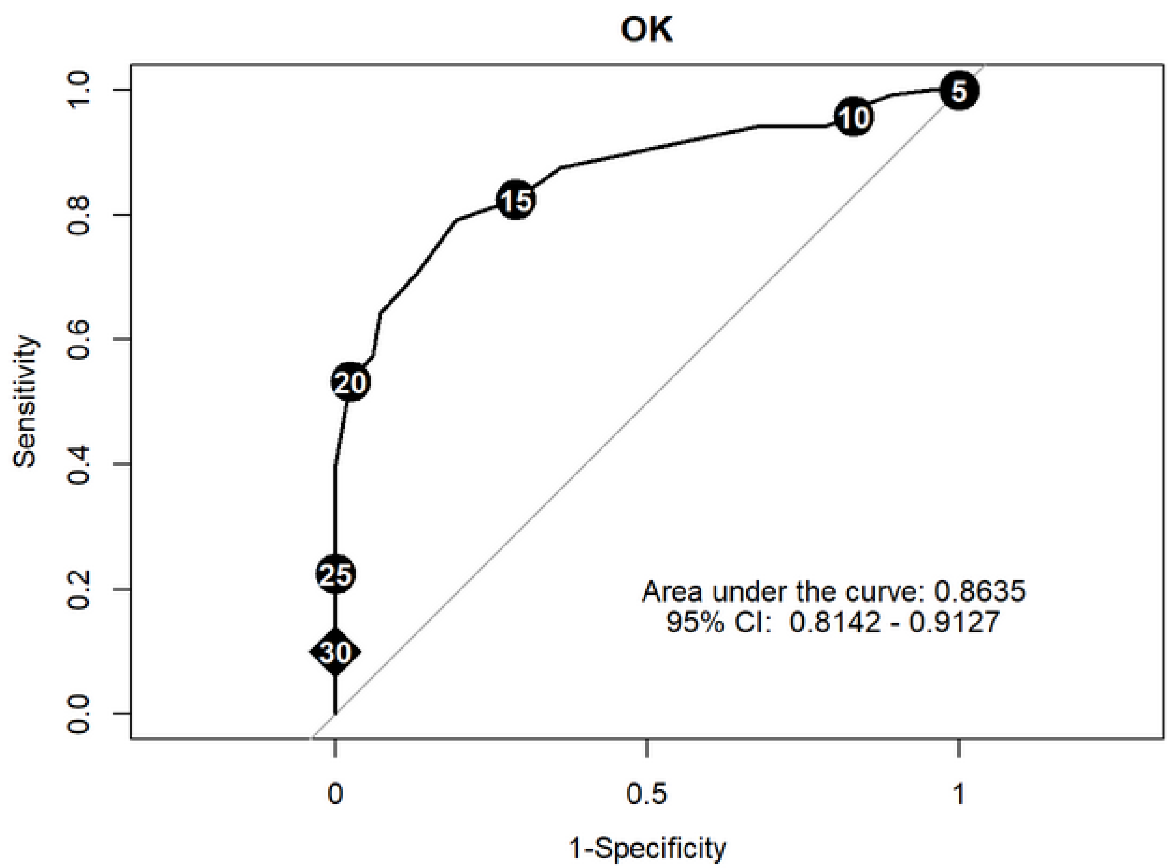

NM
Sensitivity
Specificity
Overall Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
10
20
30
40
50
Threshold

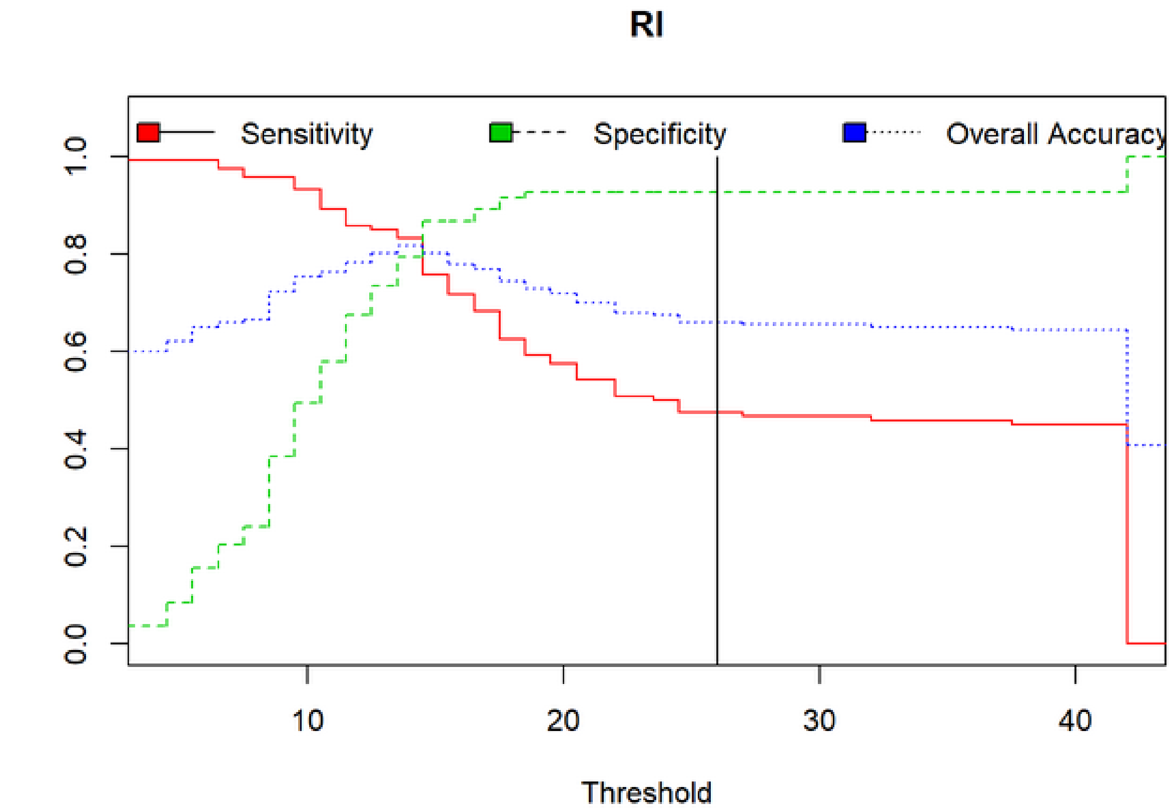

NM
Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
0
0.5
1
1-Specificity
5
10
15
20
25
30
Area under the curve: 0.7797
95% CI: 0.7139 - 0.8454

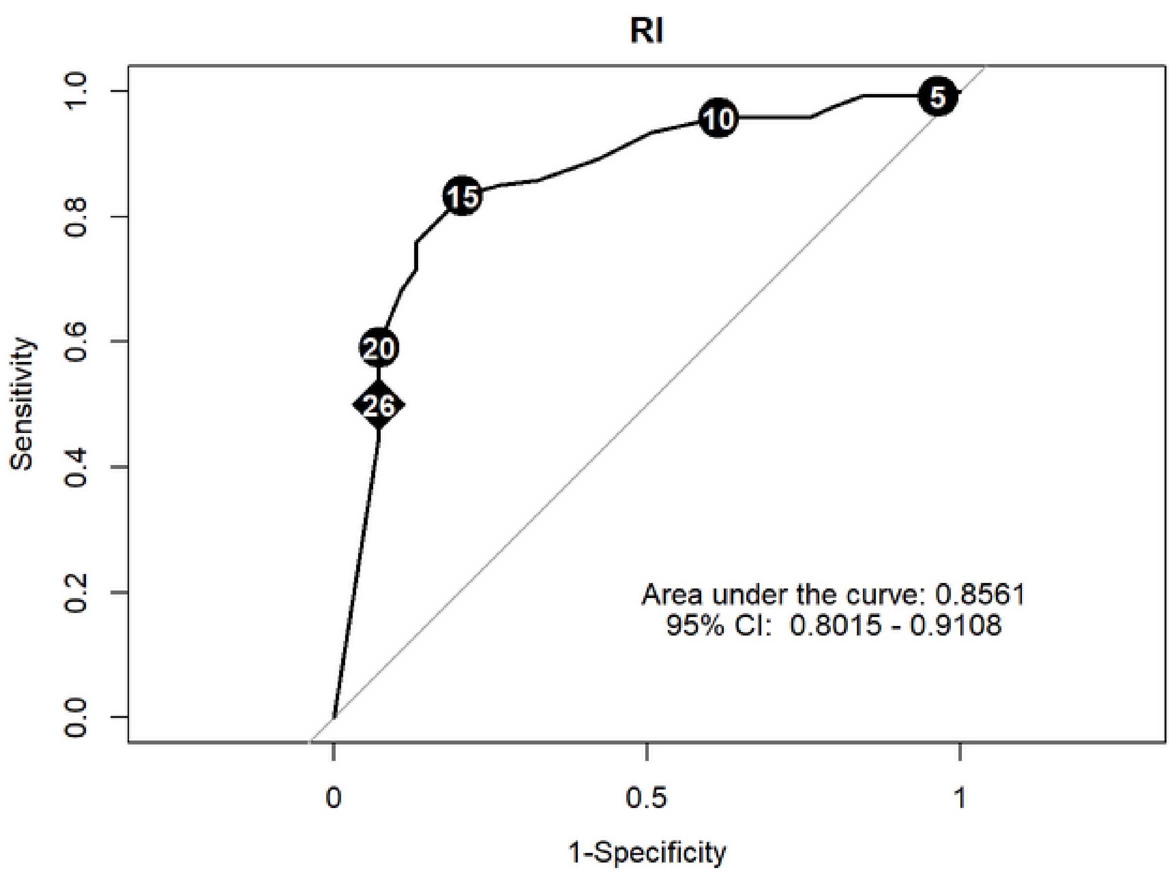

NY
Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
40
Threshold

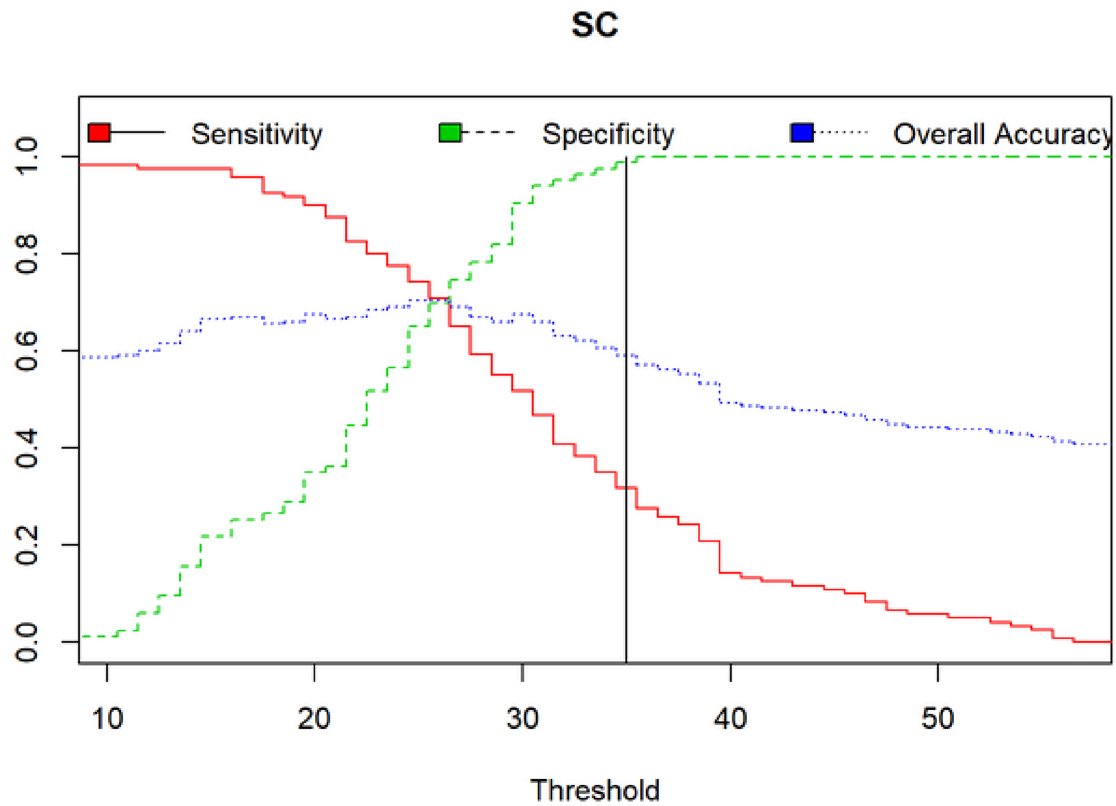

NY
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
26
30
Area under the curve: 0.8466
95% CI: 0.7911 - 0.9021
0
0.5
1
1-Specificity

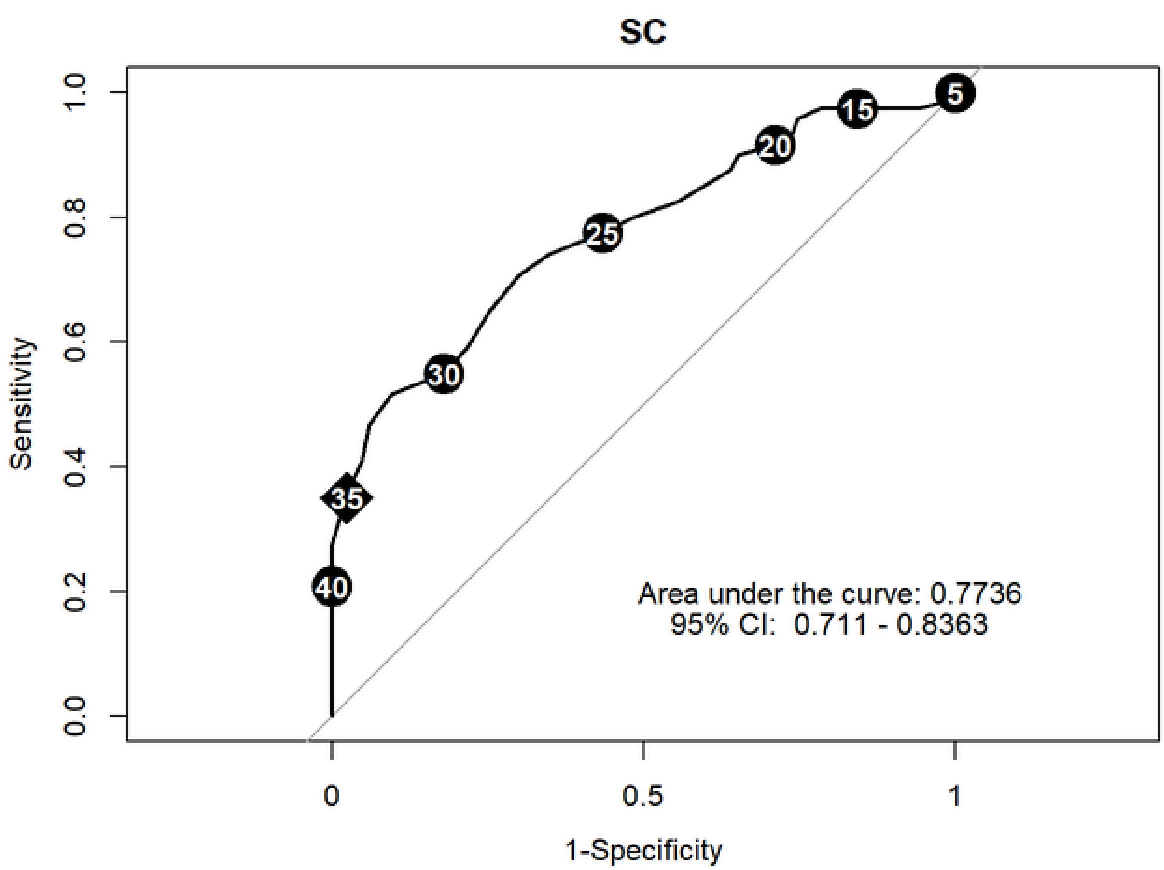

OK
Sensitivity
Specificity
Overall Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
10
20
30
40
Threshold

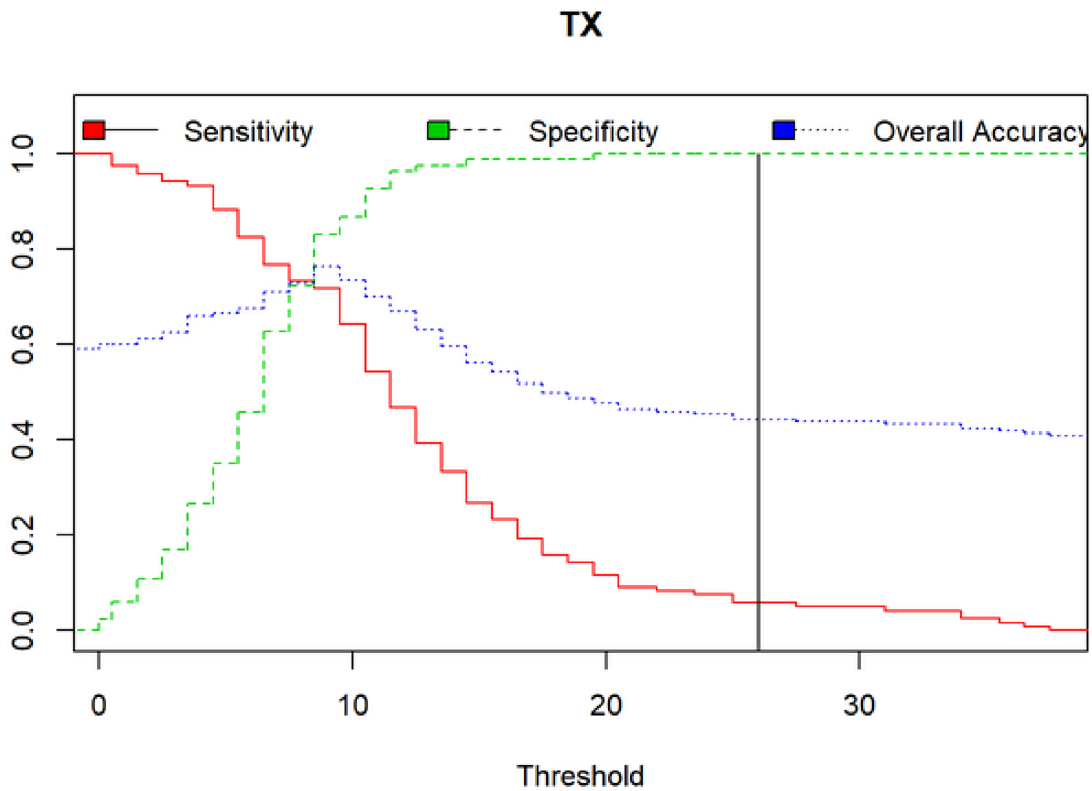

OK
Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
5
10
15
20
25
30
Area under the curve: 0.8635
95% CI: 0.8142 - 0.9127
0
0.5
1
1-Specificity

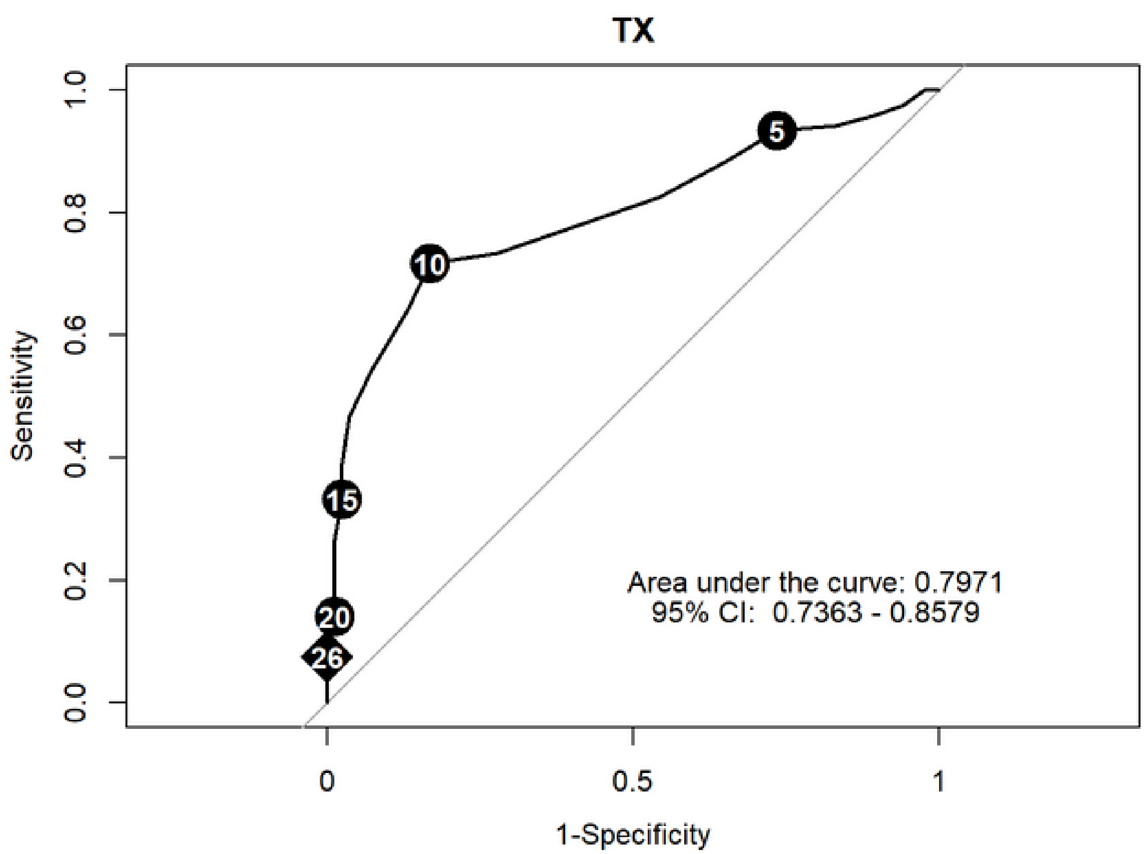

RI
Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
40
Threshold


RI
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
26
Area under the curve: 0.8561
95% CI: 0.8015 - 0.9108
0
0.5
1
1-Specificity

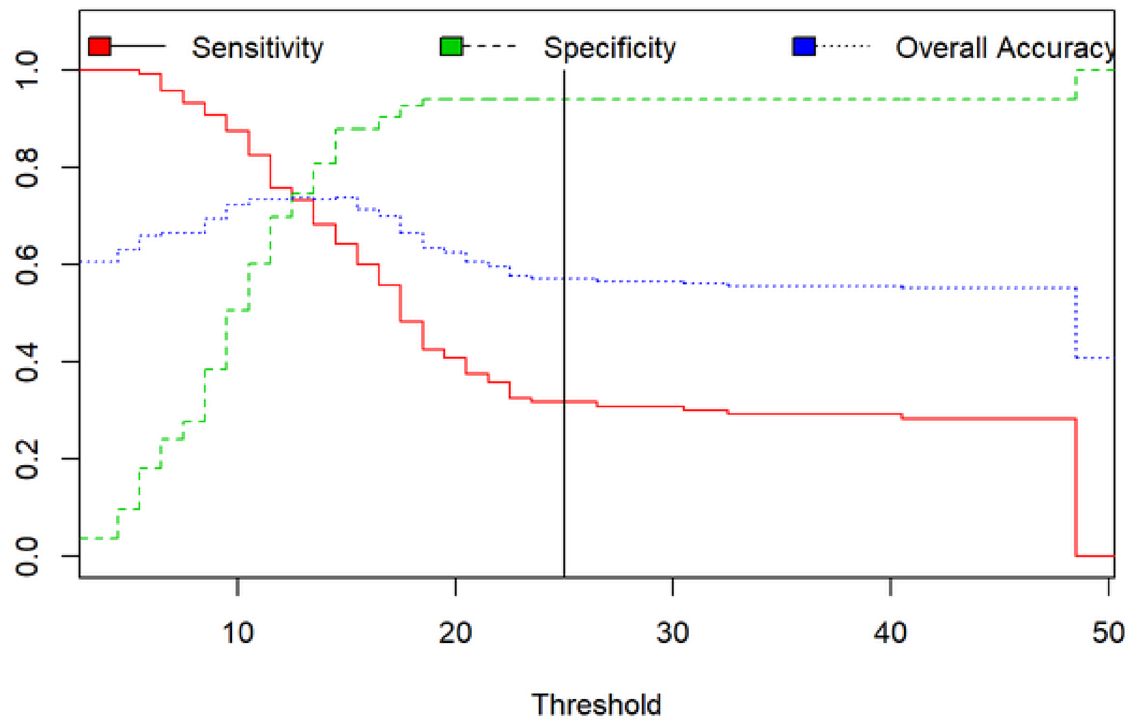

SC
Sensitivity
Specificity
Overall Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
10
20
30
40
50
Threshold


SC
Sensitivity
1-Specificity
0.0
0.2
0.4
0.6
0.8
1.0
0
0.5
1
5
15
20
25
30
35
40
Area under the curve: 0.7736
95% CI: 0.711 - 0.8363


TX
Sensitivity
Specificity
Overall Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
Threshold

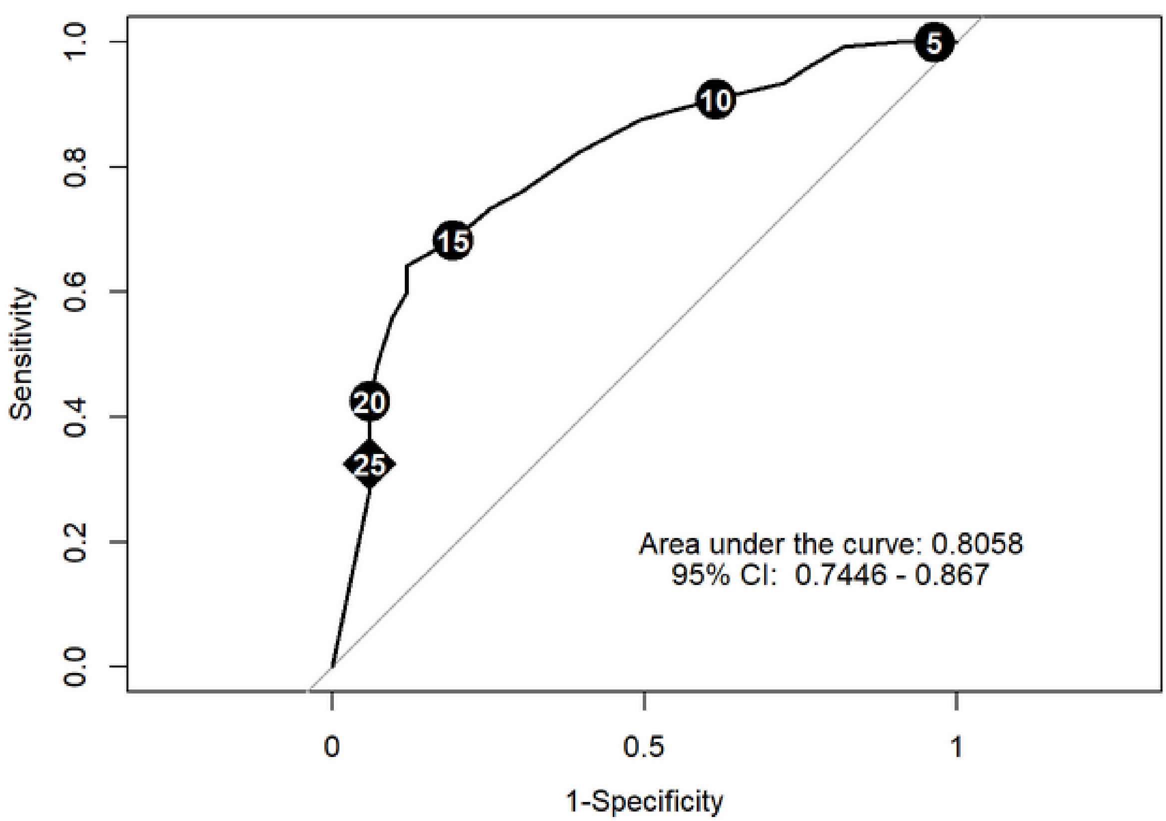

TX
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
26
Area under the curve: 0.7971
95% CI: 0.7363 - 0.8579
0
0.5
1
1-Specificity

WA

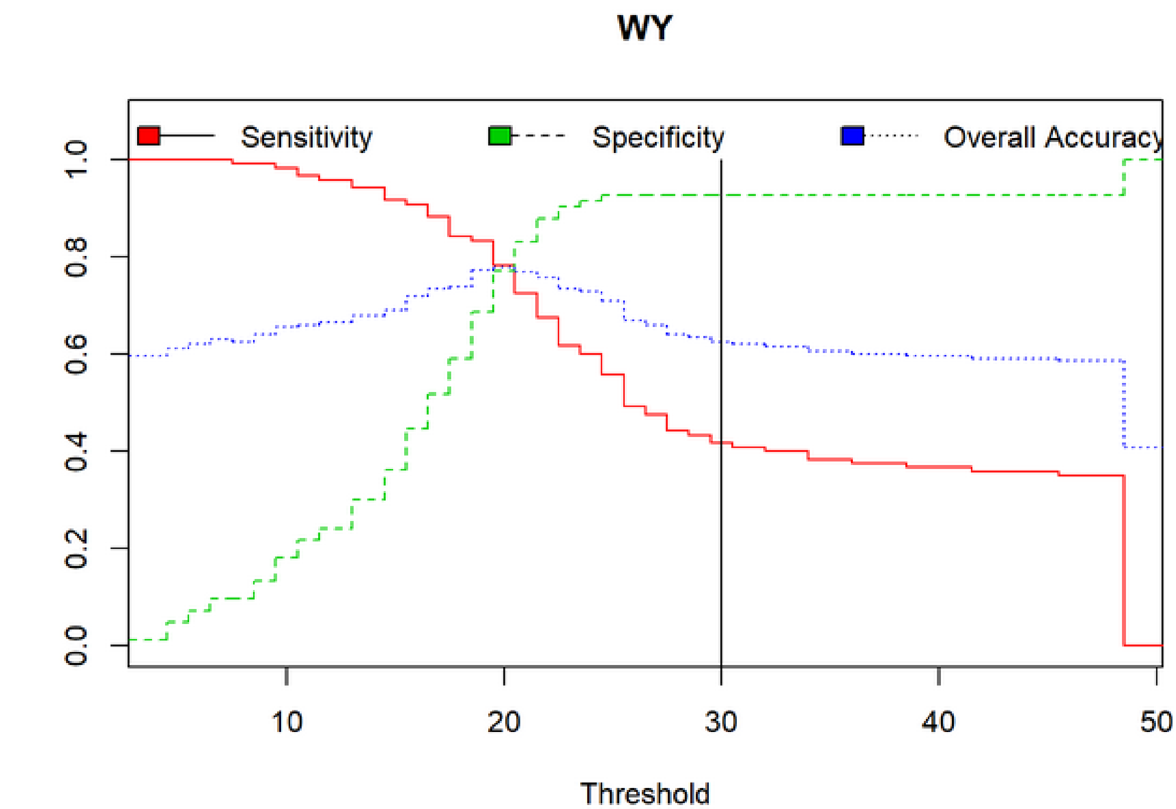


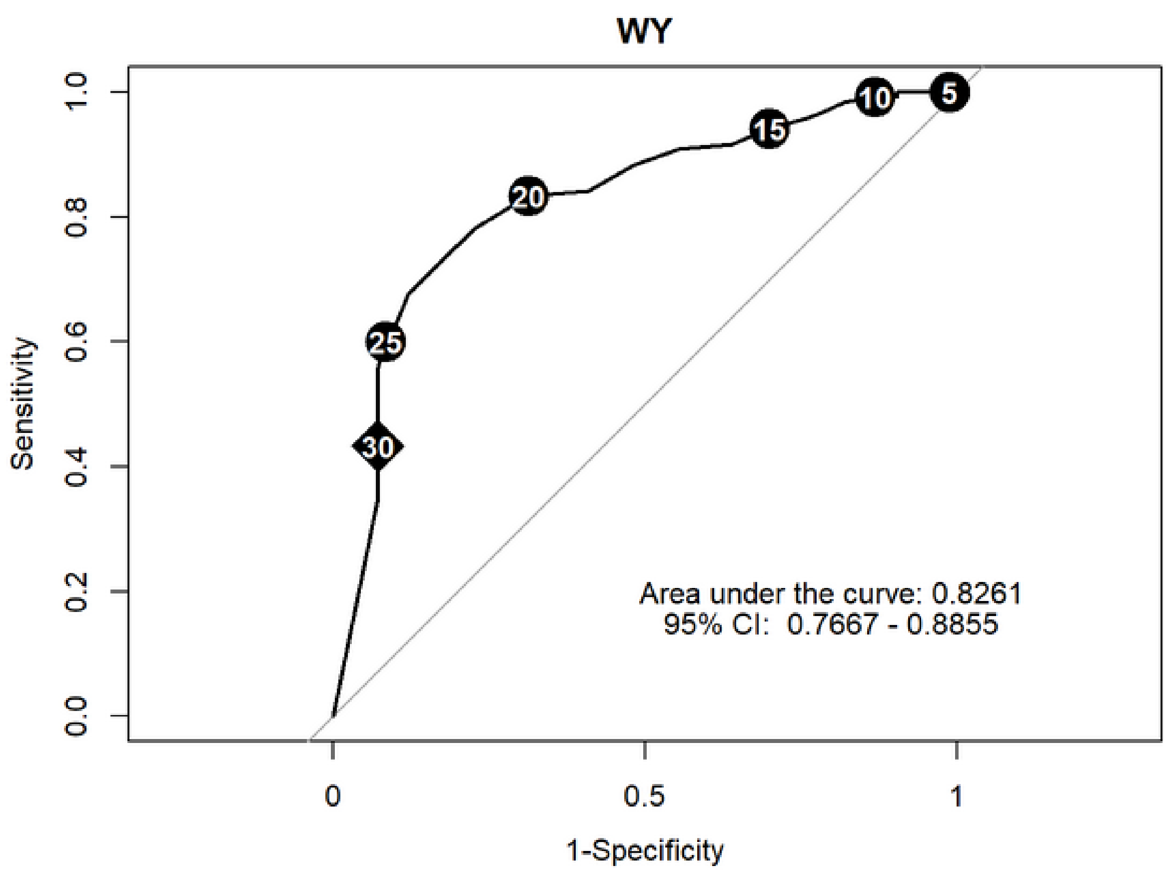
WA
Sensitivity
Specificity
Overall Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
10
20
30
40
50
Threshold

WA



WA
Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
5
10
15
20
25
Area under the curve: 0.8058
95% CI: 0.7446 - 0.867
0
0.5
1
1-Specificity


WY
Sensitivity
Specificity
Overall Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
10
20
30
40
50
Threshold


WY
Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
5
10
15
20
25
30
Area under the curve: 0.8261
95% CI: 0.7667 - 0.8855
0
0.5
1
1-Specificity

APPENDIX 4 Sensitivity/Specificity table per state-modified HLD index

Points	AK - Sens	AK - Spec	AR - Sens	AR - Spec	CA - Sens	CA - Spec	DE - Sens	DE - Spec	FL - Sens	FL - Spec	IL - Sens	IL - Spec	ME - Sens	ME - Spec	MD - Sens	MD - Spec	MA - Sens	MA - Spec	MO - Sens	MO - Spec	MT - Sens	MT - Spec	NJ - Sens	NJ - Spec	NM - Sens	NM - Spec	NY - Sens	NY - Spec	OK - Sens	OK - Spec	RI - Sens	RI - Spec	SC - Sens	SC - Spec	TX - Sens	TX - Spec	WA - Sens	WA - Spec	WY - Sens	WY - Spec	DC - Sens	DC - Spec
1	1.0000	0.0120	1.0000	0.0120	1.0000	0.0241	1.0000	0.0241	1.0000	0.0482	1.0000	0.0120	1.0000	0.0120	1.0000	0.0241	1.0000	0.0120	1.0000	0.0120	1.0000	0.0120	1.0000	0.0241	1.0000	0.0120	1.0000	0.0120	1.0000	0.0120	0.9917	0.0361	0.9833	0.0120	1.0000	0.0241	1.0000	0.0361	1.0000	0.0120	1.0000	0.0417
2	1.0000	0.0120	1.0000	0.0120	1.0000	0.0241	1.0000	0.0241	1.0000	0.1084	1.0000	0.0120	1.0000	0.0120	0.9833	0.0482	1.0000	0.0120	1.0000	0.0120	1.0000	0.0120	1.0000	0.0241	1.0000	0.0120	1.0000	0.0120	1.0000	0.0120	0.9917	0.0361	0.9833	0.0120	0.9750	0.0602	1.0000	0.0361	1.0000	0.0120	1.0000	0.0417
3	1.0000	0.0120	1.0000	0.0120	1.0000	0.0241	1.0000	0.0241	0.9833	0.1325	1.0000	0.0120	1.0000	0.0120	0.9750	0.0964	1.0000	0.0120	1.0000	0.0120	1.0000	0.0120	1.0000	0.0241	1.0000	0.0120	1.0000	0.0120	1.0000	0.0120	0.9917	0.0361	0.9833	0.0120	0.9583	0.1084	1.0000	0.0361	1.0000	0.0120	1.0000	0.0417
4	1.0000	0.0120	1.0000	0.0120	1.0000	0.0241	1.0000	0.0241	0.9500	0.1446	1.0000	0.0120	1.0000	0.0120	0.9583	0.1446	1.0000	0.0120	1.0000	0.0120	1.0000	0.0120	1.0000	0.0241	1.0000	0.0120	1.0000	0.0120	1.0000	0.0120	0.9917	0.0361	0.9833	0.0120	0.9417	0.1687	1.0000	0.0361	1.0000	0.0120	0.9697	0.0833
5	1.0000	0.0120	1.0000	0.0120	1.0000	0.0241	1.0000	0.0241	0.9417	0.2289	1.0000	0.0120	1.0000	0.0120	0.9500	0.2169	1.0000	0.0120	1.0000	0.0120	1.0000	0.0120	1.0000	0.0241	1.0000	0.0120	1.0000	0.0120	1.0000	0.0120	0.9917	0.0361	0.9833	0.0120	0.9333	0.2651	1.0000	0.0361	1.0000	0.0120	0.9545	0.1667
6	0.9917	0.0602	1.0000	0.0120	0.9917	0.0723	1.0000	0.0723	0.9083	0.3012	1.0000	0.0482	0.9917	0.0602	0.9083	0.3012	1.0000	0.0602	1.0000	0.0482	1.0000	0.0482	1.0000	0.0602	0.9917	0.0602	0.9917	0.0602	1.0000	0.0120	0.9917	0.0843	0.9833	0.0120	0.8833	0.3494	1.0000	0.0964	1.0000	0.0482	0.9394	0.2500
7	0.9833	0.1084	1.0000	0.0120	0.9833	0.1084	0.9917	0.1084	0.9000	0.4337	0.9917	0.0843	0.9833	0.1084	0.8750	0.4096	1.0000	0.0964	0.9917	0.0843	1.0000	0.0964	1.0000	0.0964	0.9833	0.1084	0.9917	0.1084	1.0000	0.0482	0.9917	0.1566	0.9833	0.0120	0.8250	0.4578	0.9917	0.1807	1.0000	0.0723	0.9091	0.3472
8	0.9667	0.1446	1.0000	0.0120	0.9583	0.1566	0.9750	0.1566	0.8667	0.4819	0.9833	0.1205	0.9667	0.1446	0.8333	0.5301	1.0000	0.1325	0.9833	0.1205	1.0000	0.1205	0.9917	0.1325	0.9667	0.1446	0.9833	0.1446	0.9917	0.1084	0.9750	0.2048	0.9833	0.0120	0.7667	0.6265	0.9583	0.2410	1.0000	0.0964	0.8636	0.4306
9	0.9500	0.1807	1.0000	0.0120	0.9417	0.2048	0.9583	0.2048	0.8500	0.5783	0.9667	0.1687	0.9500	0.1807	0.8083	0.6024	0.9917	0.1687	0.9750	0.1687	0.9917	0.1446	0.9833	0.1687	0.9500	0.1807	0.9750	0.1928	0.9750	0.1566	0.9583	0.2410	0.9833	0.0120	0.7333	0.7229	0.9333	0.2771	0.9917	0.0964	0.8182	0.5833
10	0.9333	0.2289	1.0000	0.0241	0.9250	0.2651	0.9333	0.2651	0.8250	0.6024	0.9417	0.2289	0.9333	0.2289	0.8083	0.7229	0.9917	0.2048	0.9500	0.2410	0.9917	0.2048	0.9833	0.2289	0.9333	0.2289	0.9750	0.2410	0.9583	0.1687	0.9583	0.3855	0.9833	0.0120	0.7167	0.8313	0.9083	0.3855	0.9917	0.1325	0.7727	0.6528
11	0.9167	0.3133	0.9833	0.0723	0.9083	0.3373	0.9167	0.3373	0.7917	0.6627	0.9167	0.2892	0.9167	0.3133	0.7250	0.7952	0.9917	0.2892	0.9333	0.3012	0.9917	0.2892	0.9750	0.2892	0.9083	0.3133	0.9667	0.3253	0.9417	0.2169	0.9333	0.4940	0.9833	0.0120	0.6417	0.8675	0.8750	0.5060	0.9833	0.1807	0.7424	0.7500
12	0.9000	0.4458	0.9750	0.0964	0.8833	0.4699	0.8833	0.4578	0.7750	0.7229	0.8833	0.3976	0.9000	0.4458	0.6250	0.8434	0.9833	0.4217	0.9083	0.4096	0.9750	0.4337	0.9667	0.4217	0.8917	0.4458	0.9500	0.4699	0.9417	0.3253	0.8917	0.5783	0.9833	0.0241	0.5417	0.9277	0.8250	0.6024	0.9667	0.2169	0.6212	0.7917
13	0.8667	0.4819	0.9750	0.0964	0.8417	0.5060	0.8333	0.4940	0.7417	0.7590	0.8417	0.4337	0.8667	0.4819	0.5583	0.9157	0.9583	0.4578	0.8750	0.4458	0.9500	0.4819	0.9333	0.4578	0.8500	0.4819	0.9083	0.5060	0.9083	0.4819	0.8583	0.6747	0.9750	0.0602	0.4667	0.9639	0.7583	0.6988	0.9583	0.2410	0.5152	0.8611
14	0.8583	0.5663	0.9417	0.1084	0.8000	0.6145	0.8083	0.6024	0.6917	0.8193	0.8333	0.4940	0.8583	0.5663	0.4750	0.9518	0.9500	0.5422	0.8667	0.5301	0.9500	0.5542	0.9250	0.5542	0.8417	0.5663	0.8917	0.6024	0.8750	0.6386	0.8500	0.7349	0.9750	0.0964	0.3917	0.9759	0.7333	0.7470	0.9417	0.3012	0.4848	0.9167
15	0.8083	0.6145	0.9333	0.1084	0.7583	0.6386	0.7750	0.6265	0.6333	0.8795	0.8083	0.5542	0.8083	0.6145	0.4417	0.9639	0.9167	0.5663	0.8417	0.5904	0.9250	0.5904	0.9000	0.5663	0.7917	0.6145	0.8583	0.6145	0.8250	0.7108	0.8333	0.7952	0.9750	0.1566	0.3333	0.9759	0.6833	0.8072	0.9417	0.3012	0.4394	0.9306
16	0.7667	0.6747	0.9250	0.1928	0.7250	0.6988	0.7417	0.6988	0.6167	0.8916	0.7667	0.6265	0.7667	0.6747	0.3917	0.9759	0.9000	0.6265	0.7917	0.6627	0.8750	0.6265	0.8917	0.6024	0.7417	0.6747	0.8417	0.6747	0.7917	0.8072	0.7583	0.8675	0.9750	0.2169	0.2667	0.9880	0.6417	0.8795	0.9167	0.3614	0.4091	0.9444
17	0.7500	0.7229	0.9167	0.2410	0.6917	0.7229	0.7083	0.7229	0.5750	0.8916	0.7583	0.6627	0.7500	0.7229	0.3583	0.9880	0.8833	0.6747	0.7750	0.6988	0.8583	0.6627	0.8750	0.6145	0.7250	0.7229	0.8250	0.7108	0.7083	0.8675	0.7167	0.8675	0.9583	0.2530	0.2333	0.9880	0.6000	0.8795	0.9083	0.4458	0.3636	0.9722
18	0.7333	0.7470	0.8667	0.2892	0.6667	0.7590	0.6833	0.7590	0.5583	0.9157	0.7333	0.7108	0.7333	0.7470	0.3000	0.9880	0.8833	0.7229	0.7417	0.7349	0.8500	0.6867	0.8667	0.6386	0.7083	0.7470	0.8000	0.7349	0.6417	0.9277	0.6833	0.8916	0.9583	0.2530	0.1917	0.9880	0.5583	0.9036	0.8833	0.5181	0.3182	0.9722
19	0.6833	0.7952	0.8250	0.3614	0.6333	0.8554	0.6417	0.8675	0.5333	0.9277	0.7000	0.7952	0.6833	0.7952	0.2750	0.9880	0.8250	0.7711	0.7000	0.8313	0.8000	0.7229	0.8500	0.7108	0.6583	0.7952	0.7500	0.7952	0.5750	0.9398	0.6250	0.9157	0.9250	0.2651	0.1583	0.9880	0.4833	0.9277	0.8417	0.5904	0.2576	1.0000
20	0.6250	0.8554	0.7750	0.4337	0.5917	0.8675	0.6083	0.8795	0.4500	0.9277	0.6667	0.8193	0.6250	0.8554	0.2500	0.9880	0.8083	0.8434	0.6667	0.8554	0.7833	0.8193	0.8417	0.7229	0.6000	0.8554	0.7333	0.8434	0.5333	0.9759	0.5917	0.9277	0.9167	0.2892	0.1417	0.9880	0.4250	0.9398	0.8333	0.6867	0.2273	1.0000
21	0.6083	0.8916	0.7000	0.5181	0.5750	0.9036	0.5833	0.9036	0.4333	0.9398	0.6250	0.8313	0.6083	0.8916	0.1667	1.0000	0.8000	0.8916	0.6333	0.8675	0.7833	0.8554	0.8333	0.7470	0.5750	0.8916	0.7250	0.8916	0.4667	0.9880	0.5750	0.9277	0.9000	0.3494	0.1167	1.0000	0.4083	0.9398	0.7833	0.7711	0.1970	1.0000
22	0.5750	0.8916	0.6750	0.5783	0.5417	0.9036	0.5500	0.9036	0.4000	0.9398	0.5833	0.8313	0.5750	0.8916	0.1583	1.0000	0.7750	0.8916	0.6000	0.8675	0.7667	0.8675	0.8167	0.7470	0.5417	0.8916	0.6917	0.8916	0.4000	1.0000	0.5417	0.9277	0.8750	0.3614	0.0917	1.0000	0.3750	0.9398	0.7250	0.8313	0.1515	1.0000
23	0.5417	0.9036	0.6333	0.6024	0.5167	0.9277	0.5333	0.9277	0.3583	0.9398	0.5500	0.8434	0.5417	0.9036	0.1333	1.0000	0.7500	0.9157	0.5667	0.8795	0.7333	0.8916	0.8167	0.7831	0.5083	0.9036	0.6583	0.9036	0.3333	1.0000	0.5083	0.9277	0.8250	0.4458	0.0833	1.0000	0.3583	0.9398	0.6750	0.8795	0.1212	1.0000
24	0.5000	0.9277	0.5917	0.6867	0.4667	0.9277	0.4583	0.9277	0.3417	0.9398	0.4750	0.8434	0.5000	0.9277	0.1083	1.0000	0.7500	0.9277	0.5000	0.8795	0.6917	0.9157	0.7833	0.7831	0.4417	0.9277	0.6333	0.9157	0.2750	1.0000	0.5083	0.9277	0.8000	0.5181	0.0833	1.0000	0.3250	0.9398	0.6167	0.9036	0.0909	1.0000
25	0.4583	0.9277	0.5167	0.7952	0.4250	0.9277	0.4167	0.9277	0.3333	0.9398	0.4083	0.8434	0.4583	0.9277	0.1083	1.0000	0.7167	0.9277	0.4583	0.8795	0.6667	0.9398	0.7750	0.7831	0.3917	0.9277	0.6083	0.9157	0.2250	1.0000	0.5000	0.9277	0.7750	0.5663	0.0750	1.0000	0.3167	0.9398	0.6000	0.9157	0.0758	1.0000
26	0.4250	0.9398	0.4500	0.8675	0.4083	0.9398	0.4000	0.9398	0.3250	0.9398	0.3917	0.8554	0.4250	0.9398	0.0917	1.0000	0.6917	0.9277	0.4333	0.8916	0.6167	0.9639	0.7667	0.7952	0.3583	0.9398	0.5917	0.9277	0.1667	1.0000	0.4750	0.9277	0.7417	0.6506	0.0583	1.0000	0.3167	0.9398	0.5583	0.9277	0.0758	1.0000
27	0.4000	0.9398	0.3667	0.8916	0.3583	0.9398	0.3583	0.9398	0.3083	0.9398	0.3750	0.8675	0.4000	0.9398	0.0917	1.0000	0.6833	0.9277	0.4250	0.9036	0.5667	0.9639	0.7333	0.7952	0.3250	0.9398	0.5583	0.9277	0.1583	1.0000	0.4750	0.9277	0.7083	0.6988	0.0583	1.0000	0.3167	0.9398	0.4917	0.9277	0.0606	1.0000
28	0.3500	0.9398	0.3250	0.9518	0.3083	0.9398	0.3167	0.9398	0.3000	0.9398	0.3333	0.8795	0.3500	0.9398	0.0833	1.0000	0.6333	0.9277	0.4083	0.9157	0.5083	0.9639	0.6833	0.7952	0.2750	0.9398	0.5083	0.9277	0.1250	1.0000	0.4667	0.9277	0.6500	0.7470	0.0583	1.0000	0.3083	0.9398	0.4750	0.9277	0.0606	1.0000
29	0.3417	0.9398	0.3167	0.9639	0.3083	0.9398	0.3083	0.9398	0.2917	0.9398	0.3250	0.8795	0.3417	0.9398	0.0833	1.0000	0.6250	0.9277	0.3917	0.9157	0.4500	0.9639	0.6833	0.7952	0.2583	0.9398	0.5000	0.9277	0.1167	1.0000	0.4667	0.9277	0.5917	0.7831	0.0500	1.0000	0.3083	0.9398	0.4417	0.9277	0.0606	1.0000
30	0.3250	0.9398	0.2750	1.0000	0.3000	0.9398	0.3083	0.9398	0.2917	0.9398	0.3167	0.8795	0.3250	0.9398	0.0750	1.0000	0.6167	0.9277	0.3917	0.9157	0.4333	0.9639	0.6833	0.7952	0.2417	0.9398	0.4917	0.9277	0.1000	1.0000	0.4667	0.9277	0.5500	0.8193	0.0500	1.0000	0.3083	0.9398	0.4333	0.9277	0.0303	1.0000
31	0.3167	0.9398	0.2083	1.0000	0.3000	0.9398	0.2917	0.9398	0.2917	0.9398	0.2917	0.8795	0.3167	0.9398	0.0667	1.0000	0.6083	0.9277	0.3833	0.9157	0.4083	0.9639	0.6833	0.7952	0.2333	0.9398	0.4833	0.9277	0.0917	1.0000	0.4667	0.9277	0.5167	0.9036	0.0500	1.0000	0.3083	0.9398	0.4167	0.9277	0.0303	1.0000
32	0.3167	0.9398	0.1917	1.0000	0.3000	0.9398	0.2917	0.9398	0.2833	0.9398	0.2750	0.8795	0.3167	0.9398	0.0500	1.0000	0.6000	0.9277	0.3667	0.9157	0.3833	0.9639	0.6750	0.7952	0.2167	0.9398	0.4833	0.9277	0.0917	1.0000	0.4667	0.9277	0.4667	0.9398	0.0417	1.0000	0.3000	0.9398	0.4083	0.9277	0.0303	1.0000
33	0.3083	0.9398	0.1583	1.0000	0.2917	0.9398	0.2917	0.9398	0.2833	0.9398	0.2667	0.8795	0.3083	0.9398	0.0500	1.0000	0.5917	0.9277	0.3583	0.9157	0.3667	0.9639	0.6750	0.7952	0.2083	0.9398	0.4750	0.9277	0.0833	1.0000	0.4583	0.9277	0.4083	0.9518	0.0417	1.0000	0.3000	0.9398	0.4000	0.9277	0.0152	1.0000
34	0.3000	0.9398	0.1250	1.0000	0.2833	0.9398	0.2833	0.9398	0.2833	0.9398	0.2667	0.8795	0.3000	0.9398	0.0417	1.0000	0.5750	0.9277	0.3583	0.9157	0.3667	0.9639	0.6750	0.7952	0.2000	0.9398	0.4583	0.9277	0.0750	1.0000	0.4583	0.9277	0.3833	0.9639	0.0417	1.0000	0.2917	0.9398	0.4000	0.9277	0.0152	1.0000
35	0.2833	0.9398	0.1000	1.0000	0.2833	0.9398	0.2833	0.9398	0.2833	0.9398	0.2667	0.8795	0.2833	0.9398	0.0417	1.0000	0.5667	0.9277	0.3583	0.9157	0.3250	0.9639	0.6750	0.7952	0.1833	0.9398	0.4583	0.9277	0.0667	1.0000	0.4583	0.9277	0.3500	0.9759	0.0250	1.0000	0.2917	0.9398	0.3833	0.9277	0.0152	1.0000
36	0.2833	0.9398	0.0833	1.0000	0.2833	0.9398	0.2833	0.9398	0.2750	0.9398	0.2667	0.8795	0.2833	0.9398	0.0417	1.0000	0.5667	0.9277	0.3583	0.9157	0.3083	0.9639	0.6750	0.7952	0.1833	0.9398	0.4583	0.9277	0.0583	1.0000	0.4583	0.9277	0.3167	0.9880	0.0250	1.0000	0.2917	0.9398	0.3833	0.9277	0.0152	1.0000
37	0.2833	0.9398	0.0583	1.0000	0.2750	0.9398	0.2750	0.9398	0.2750	0.9398	0.2583	0.8795	0.2833	0.9398	0.0333	1.0000	0.5583	0.9277	0.3583	0.9157	0.2750	0.9639	0.6750	0.7952	0.1833	0.9398	0.4500	0.9277	0.0500	1.0000	0.4583	0.9277	0.2750	1.0000	0.0167	1.0000	0.2917	0.9398	0.3750	0.9277	0.0152	1.0000

38	0.2833	0.9398	0.0583	1.0000	0.2750	0.9398	0.2750	0.9398	0.2750	0.9398	0.2583	0.8795	0.2833	0.9398	0.0333	1.0000	0.5583	0.9277	0.3583	0.9157	0.2750	0.9639	0.6750	0.7952	0.1750	0.9398	0.4500	0.9277	0.0500	1.0000	0.4583	0.9277	0.2583	1.0000	0.0083	1.0000	0.2917	0.9398	0.3750	0.9277	0.0152	1.0000
39	0.2750	0.9398	0.0500	1.0000	0.2750	0.9398	0.2750	0.9398	0.2750	0.9398	0.2583	0.8795	0.2750	0.9398	0.0167	1.0000	0.5500	0.9277	0.3583	0.9157	0.2667	0.9639	0.6750	0.7952	0.1750	0.9398	0.4500	0.9277	0.0417	1.0000	0.4500	0.9277	0.2417	1.0000	0.0000	1.0000	0.2917	0.9398	0.3750	0.9277	0.0000	1.0000
40	0.2750	0.9398	0.0500	1.0000	0.2750	0.9398	0.2750	0.9398	0.2667	0.9398	0.2500	0.8795	0.2750	0.9398	0.0083	1.0000	0.5500	0.9277	0.3500	0.9157	0.2667	0.9639	0.6667	0.7952	0.1750	0.9398	0.4500	0.9277	0.0417	1.0000	0.4500	0.9277	0.2083	1.0000	0.0000	1.0000	0.2917	0.9398	0.3667	0.9277	0.0000	1.0000
41	0.2750	0.9398	0.0333	1.0000	0.2750	0.9398	0.2750	0.9398	0.2667	0.9398	0.2500	0.8795	0.2750	0.9398	0.0083	1.0000	0.5417	0.9277	0.3500	0.9157	0.2667	0.9639	0.6667	0.7952	0.1667	0.9398	0.4500	0.9277	0.0333	1.0000	0.4500	0.9277	0.1417	1.0000	0.0000	1.0000	0.2917	0.9398	0.3667	0.9277	0.0000	1.0000
42	0.2750	0.9398	0.0333	1.0000	0.2667	0.9398	0.2750	0.9398	0.2667	0.9398	0.2500	0.8795	0.2750	0.9398	0.0000	1.0000	0.5417	0.9277	0.3500	0.9157	0.2667	0.9639	0.6667	0.7952	0.1667	0.9398	0.4417	0.9277	0.0333	1.0000	0.4500	0.9277	0.1333	1.0000	0.0000	1.0000	0.2833	0.9398	0.3667	0.9277	0.0000	1.0000
43	0.2667	0.9398	0.0250	1.0000	0.2667	0.9398	0.2750	0.9398	0.2667	0.9398	0.2417	0.8795	0.2667	0.9398	0.0000	1.0000	0.5417	0.9277	0.3500	0.9157	0.2500	0.9639	0.6667	0.7952	0.1583	0.9398	0.4417	0.9277	0.0333	1.0000	0.0000	1.0000	0.1250	1.0000	0.0000	1.0000	0.2833	0.9398	0.3583	0.9277	0.0000	1.0000
44	0.2667	0.9398	0.0250	1.0000	0.2667	0.9398	0.2667	0.9398	0.2667	0.9398	0.2417	0.8795	0.2667	0.9398	0.0000	1.0000	0.5417	0.9277	0.3500	0.9157	0.2500	0.9639	0.6667	0.7952	0.1583	0.9398	0.4417	0.9277	0.0250	1.0000	0.0000	1.0000	0.1167	1.0000	0.0000	1.0000	0.2833	0.9398	0.3583	0.9277	0.0000	1.0000
45	0.2667	0.9398	0.0167	1.0000	0.2667	0.9398	0.2667	0.9398	0.0000	1.0000	0.2417	0.8795	0.2667	0.9398	0.0000	1.0000	0.5417	0.9277	0.3500	0.9157	0.2500	0.9639	0.6667	0.7952	0.1583	0.9398	0.4417	0.9277	0.0250	1.0000	0.0000	1.0000	0.1167	1.0000	0.0000	1.0000	0.2833	0.9398	0.3583	0.9277	0.0000	1.0000
46	0.2667	0.9398	0.0083	1.0000	0.2667	0.9398	0.2667	0.9398	0.0000	1.0000	0.2417	0.8795	0.2667	0.9398	0.0000	1.0000	0.5333	0.9277	0.3500	0.9157	0.2417	0.9639	0.6667	0.7952	0.1583	0.9398	0.4417	0.9277	0.0000	1.0000	0.0000	1.0000	0.1083	1.0000	0.0000	1.0000	0.2833	0.9398	0.3583	0.9277	0.0000	1.0000
47	0.0000	1.0000	0.0083	1.0000	0.0000	1.0000	0.2667	0.9398	0.0000	1.0000	0.2333	0.8795	0.0000	1.0000	0.0000	1.0000	0.5333	0.9277	0.3417	0.9157	0.2333	0.9639	0.0000	1.0000	0.1500	0.9398	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.1000	1.0000	0.0000	1.0000	0.2833	0.9398	0.3500	0.9277	0.0000	1.0000
48	0.0000	1.0000	0.0083	1.0000	0.0000	1.0000	0.2667	0.9398	0.0000	1.0000	0.2333	0.8795	0.0000	1.0000	0.0000	1.0000	0.5333	0.9277	0.3417	0.9157	0.2333	0.9639	0.0000	1.0000	0.1500	0.9398	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0833	1.0000	0.0000	1.0000	0.2833	0.9398	0.3500	0.9277	0.0000	1.0000
49	0.0000	1.0000	0.0083	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.2250	0.8795	0.0000	1.0000	0.0000	1.0000	0.5333	0.9277	0.3333	0.9157	0.2333	0.9639	0.0000	1.0000	0.1500	0.9398	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0667	1.0000	0.0000	1.0000	0.2833	0.9398	0.3500	0.9277	0.0000	1.0000
50	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.5333	0.9277	0.0000	1.0000	0.2250	0.9639	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0583	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
51	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.5333	0.9277	0.0000	1.0000	0.2250	0.9639	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0583	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
52	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.5333	0.9277	0.0000	1.0000	0.2250	0.9639	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0500	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
53	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.5333	0.9277	0.0000	1.0000	0.2250	0.9639	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0500	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
54	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.5333	0.9277	0.0000	1.0000	0.2250	0.9639	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0417	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
55	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.5333	0.9277	0.0000	1.0000	0.2083	0.9639	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0333	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
56	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.5333	0.9277	0.0000	1.0000	0.2083	0.9639	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0250	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
57	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.5333	0.9277	0.0000	1.0000	0.2083	0.9639	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0083	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
58	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.5333	0.9277	0.0000	1.0000	0.2083	0.9639	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
59	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
60	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
61	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000

APPENDIX 3: Eligibility requirements per HLD-using states

	AK	AR	CA	DE	FL	IL	ME	MD	MA	MO	MT	NJ	NM	NY	OK	RI	SC	TX	WA	WY
Minimum Age	-	13[1]	13[1]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	13`	-	12
Maximum Age	21	21	21	21	21	21	21	21	21	21	21	21	21	21	18[4]	21	-	21	21	19
No deciduous teeth	-	✓[2, 3]	✓[2]	✓	-	✓	-	-	-	✓[2, 3]	-	✓	-	✓[2, 3]	✓[1]	-	-	✓[2, 3]	✓	-

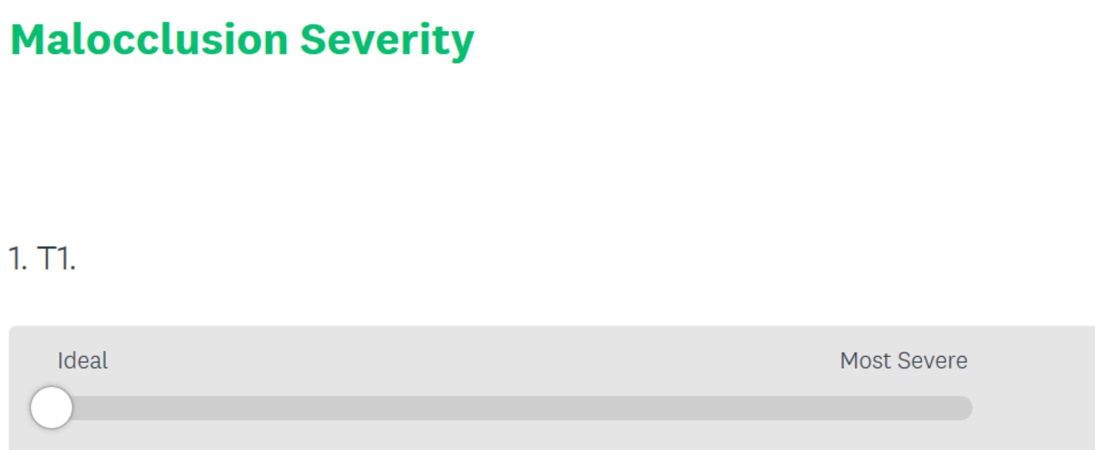
[1] Cleft and craniofacial conditions exempted from age requirement

[2] Over-retained deciduous teeth permitted secondary to oligodontia or ectopia

[3] Deciduous teeth with no roots remaining permitted

[4] At time of application

APPENDIX 4: Digital visual analog scale from SurveyMonkey® used by the expert panel



APPENDIX 5: Expert Panel pre-evaluation oral prompt

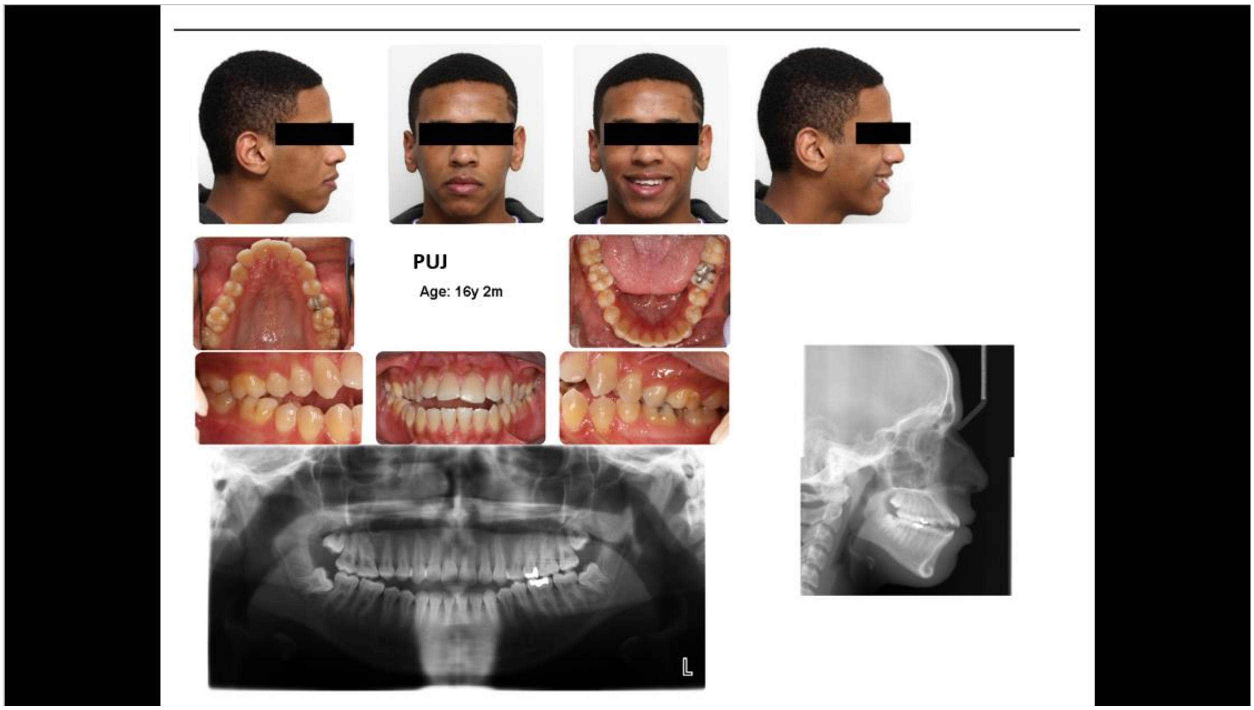
As an orthodontic specialist, you see a wide range of malocclusions come through your offices on a daily basis. In this exercise, you are an orthodontic consultant tasked with reviewing the orthodontic records of patients seeking orthodontic treatment. Pre-treatment extra- and intra-oral photographs as well as panoramic and cephalometric radiographs will be presented to you for evaluation. In addition, you will be given a digital 100-point visual analog scale (VAS). Please review the records and rate the severity of the patient's malocclusion on a scale of 0 to 100, with 0 being a perfect or ideal occlusion and 100 being the most severe malocclusion imaginable. To do so, simply click on the VAS. Please do not drag the slider.

Please keep the following in mind as you evaluate the patient records:

1) Please note that you are rating for malocclusion severity and NOT treatment difficulty.
2) Please be sure to review all records presented to you.
3) Please be aware that you may recognize or have treated some patients since they were all screened for orthodontic treatment at the University of Washington. Please make every attempt to disassociate your personal experiences with the patient from your evaluation of their initial malocclusion severity.

You will have 20 seconds to evaluate each patient record layout which will include baseline extra- and intra-oral photographs as well as panoramic and cephalometric radiographs. A coffee/restroom break will be given at the halfway point. To help you get a sense for the pace and format of the rating session, a series of 10 training cases will be shown prior to the start of the evaluation. As you rate each malocclusion, please keep in mind that at the end of the rating session, that you will be asked to provide an indicated treatment need point above which you consider orthodontic treatment to be essential.

APPENDIX 6 – Example of a pre-treatment record slide



APPENDIX 7 – Handicapping Labiolingual Deviations Washington State Modification as of June 30, 2018

Washington State Health Care Authority

Orthodontic Information

MEDICAID AUTHORIZATIONS – ORTHO
PO Box 45535
Olympia, WA 98504-5535

All blank fields below must be completed; please see example form

Provider name	Patient's name: Last	First
Billing provider number	Performing provider number	
Client ID	Client birth date	Clie

PART I. Orthodontic treatment requested (check box below) and diagnostic information

☐ Case study only ☐ Fixed appliance therapy ☐ Limited transitic
☐ Interceptive treatment ☐ Comprehensive full treatment ☐ Transfer case (if required to con

Tentative treatment plan:

Functional concerns:

Will the client require orthognathic surgery? ☐ Yes ☐ No

Has the client seen a general dentist in the last 12 months? ☐ Yes ☐ No

Brief in

Stage of dentition: ☐ Primary ☐ Adolescent ☐ Mixed/Transitional

Client's chief complaint

Anterior teeth:

Overjet		mm
Overbite		mm
Openbite		mm
Midline		mm

Crossbite:
Indicate teeth involved

Habits

Posterior teeth:

Angle Classification:
Skeletal classification: (check one)
☐ Class I ☐ Class II ☐ Class III

Musculature: tone and function

Dental classification: (check one)

Right	☐ Class I	☐ E to E	☐ Class II	☐ Class III
Left	☐ Class I	☐ E to E	☐ Class II	☐ Class III

Symmetry of arches

Crossbite:
Indicate teeth involved

Anterior
Crowding (Approximate) Spacing

Temporomandibular dysfunctio

Crowding			Spacing		
MAX		mm	MAX		mm
MAND		mm	MAND		mm

HCA 13-666 (7/18)

Missing teeth (list) Oral hygiene: ☐ Good ☐

	Yes	Tooth/location
Ectopic eruption (Numbers of teeth excluding third molar(s):	☐	N/A
Missing (indicate teeth):	☐	N/A
Impacted (indicate teeth):	☐	N/A
Ankylosed (indicate teeth):	☐	N/A
Supernumerary (indicate location):	☐	N/A

Restoration or caries problems
N/A

Other medical or dental problems:
N/A

PART II. Overbite, crossbite or overjet information. See instructions for further information.

Place an "x" for each condition that applies

1. Client has a deep impinging overbite when lower incisors are destroying the soft tissue of the palate. Ph confirms deep impinging overbites when the lower incisors are destroying the soft tissue of the palate r
2. Client has a crossbite of individual anterior teeth when destruction of the soft tissue is present. Recessi due to crossbite must be more than 1mm and the recession must not be due to the lower crowding but anterior crossbite.
3. Client has an overjet greater than 9mm with incompetent lips or reverse overjet greater than 3.5mm wi and speech difficulties. If this is applicable, provide a color photo using either a probe or ruler to demon
4. Client has a negative overjet relative to a skeletal Class III. A recent cephalometric radiographic image n confirm this condition.

PART III. Handicapping Labiolingual Deviation Index (HLD). See instructions regarding scoring.

1. Overjet in mm.
2. Overbite in mm.
3. Mandibular protrusion. X 5 =
4. Openbite in mm. 0 X 4 =

If both anterior crowding and ectopic eruptions are present in the anterior portion of the mouth, score onl Do not score both conditions.

5. Ectopic eruption: Count each tooth, excluding third molars X 3 =
6. Anterior crowding: Anterior arch length insufficiency must exceed 3.5mm; score one point for maxilla a 2 points maximum for anterior crowding. The maximum number of points for this item is therefore 10 p lower). X 5 =
7. Posterior unilateral crossbite: This condition involves two or more adjacent teeth, one of which must be must be one in which the maxillary posterior teeth involved may be both palatal or both completely buc mandibular posterior teeth. The presence of posterior unilateral crossbite is indicated by a score of 4 on left and right posterior crossbite are present, score 4 for each side.

PROVIDER'S ESTIMATED TOTAL HLD SCORE (REQUIRED)

PLEASE NOTE: The HLD scoring is a guideline for your use and reference, and rescoring may be completed by ou send all required information referred to in Billing Instruction and WAC. The department will make the final dec scoring. This information may not be used to predetermine coverage in order to charge the client.

Examination completed by: Print name

I certify that I am the performing provider and that the medical necessity information is true, accurate, an knowledge. I understand that any falsification, omission, or concealment of material fact in those sections liability.

PERFORMING PROVIDER SIGNATURE	Print name (INCLUDE CREDEN