Standardization Before Prediction

Establishing Minimum Data Elements for Osteoradionecrosis as a Foundation for Predictive Modeling

Abdallah S.R. Mohamed, MD, PhD
Assistant Professor, Department of Radiation Oncology
Baylor College of Medicine | MD Anderson Cancer Center
The International ORAL Consortium

NIDCR Data Science – AI/ML in Diagnostics and Treatment
2026 IADR/AADOCR/CADR General Session & Exhibition

Disclosures & Funding

Funding Support:

NIH/NIDCR: U01DE032168, R01DE034780

NIH/NCI: R01CA258827, P01CA285249, P01CA278716

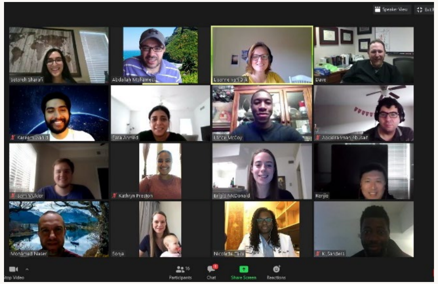
Relevant Relationships:

ORAL Consortium Member | Co-Investigator on 5 active NIH grants | PREDMORN Consortium Member

Acknowledgments

THE UNIVERSITY OF TEXAS
MD Anderson Cancer Center
Making Cancer History®

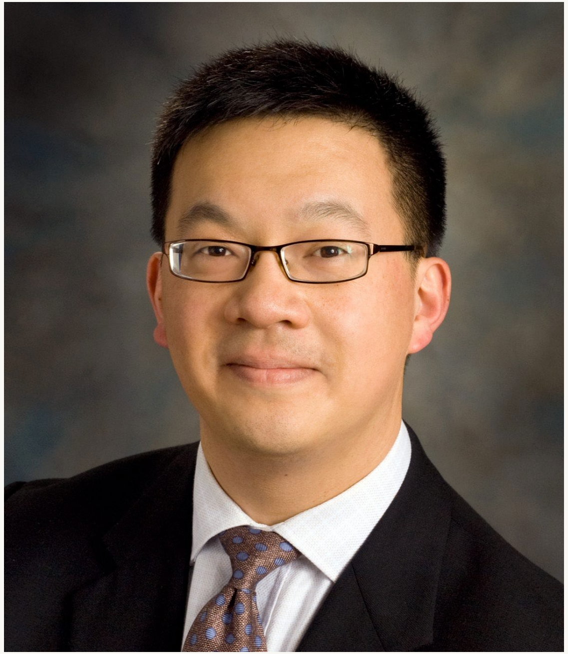
Dr. Fuller and Dr. Moreno Lab

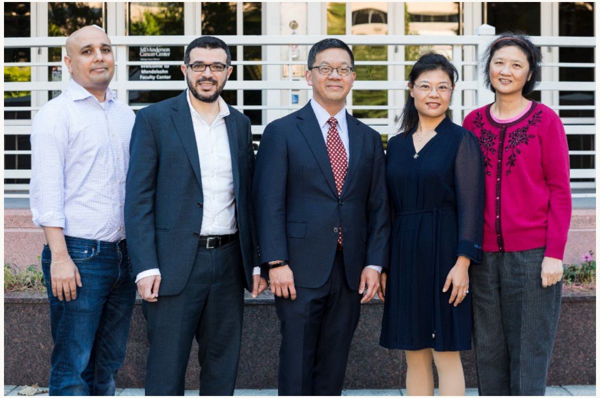


Acknowledgments

THE UNIVERSITY OF TEXAS
MD Anderson ~~Cancer~~ Center
Making Cancer History®

Dr. Stephen Y. Lai and the Lai Lab





Agenda

I The Problem: Challenges of ORN Prediction

II The Solution: ORAL Consortium & Minimum Data Elements

III The Evidence: AI/ML for ORN — What Works & What Breaks

IV The Path Forward: CDEs as Infrastructure for Next-Gen Models

Osteoradionecrosis of the Jaw (ORNJ)

A severe iatrogenic disease: bone death after radiation therapy to the head and neck

5–15%

Incidence in HNC survivors

$170K

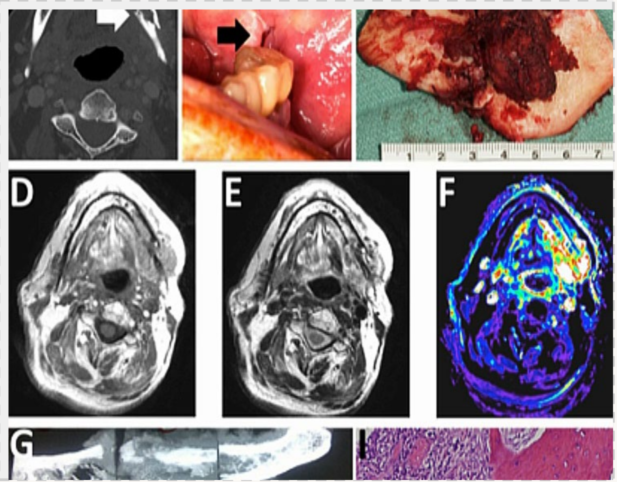
Cost per patient for advanced cases

~2 yrs

Avg. onset post-RT; progressive if untreated

No ICD

No specific code until 2023



Risk factors: poor oral hygiene, pre/post-RT dental extractions, high mandibular radiation dose-volumes, smoking, periodontal disease

Moreno AC et al. IJROBP 2025; Watson EE et al. JCO 2024; Peterson DE et al. JCO 2024

The Definitional Chaos

9+ Published definitions of ORNJ

16+ Classification / staging systems over 4 decades

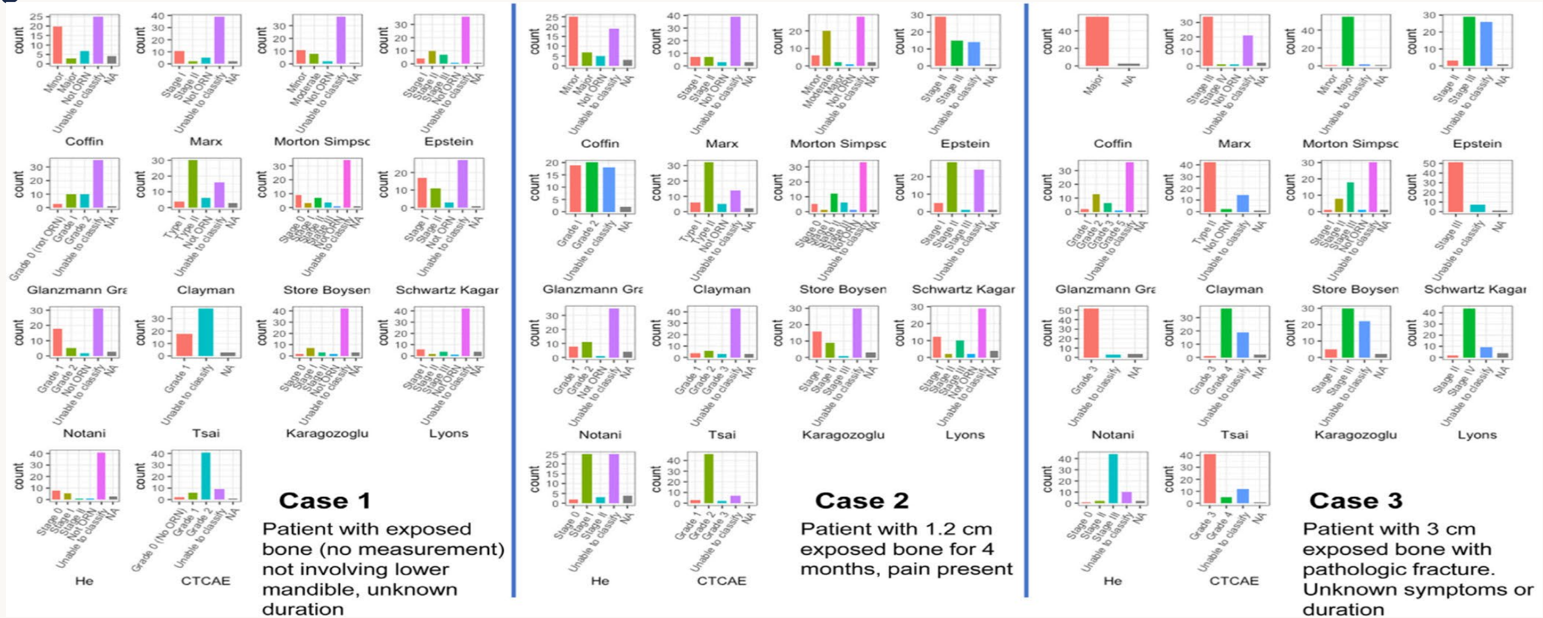
76% Inability to classify cases using some existing systems

Consequences for AI/ML

- Inaccurate incidence estimation
- Reporting ambiguity & underdiagnosis
- Poor cross-study comparability
- Inconsistent training labels for models
- Blocks multi-institutional validation

Moreno AC et al. IJROBP 2025; 122(2):341–354

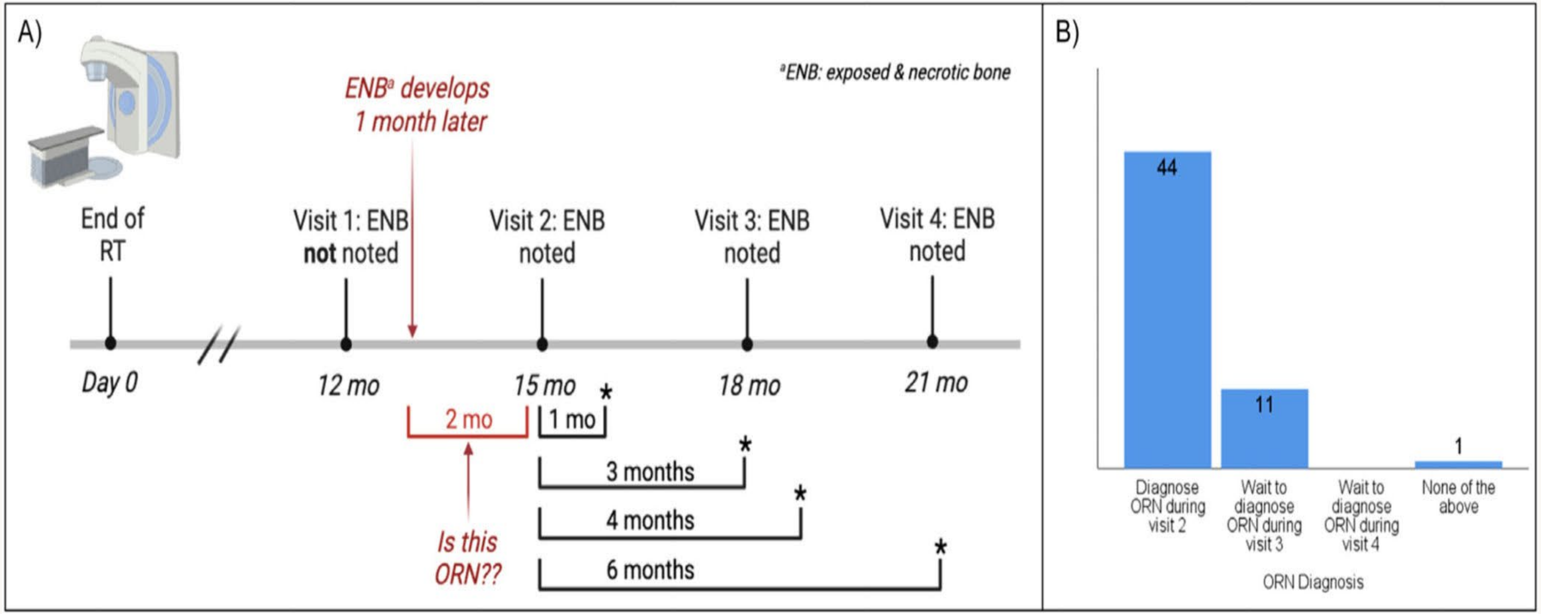
Expert Classification Using 14 Existing Systems



3 descriptive ORN cases classified by 14 systems. Magenta bars = 'unable to classify.' Note the dramatic variability across systems for the same patient.

Moreno AC et al. IJROBP 2025

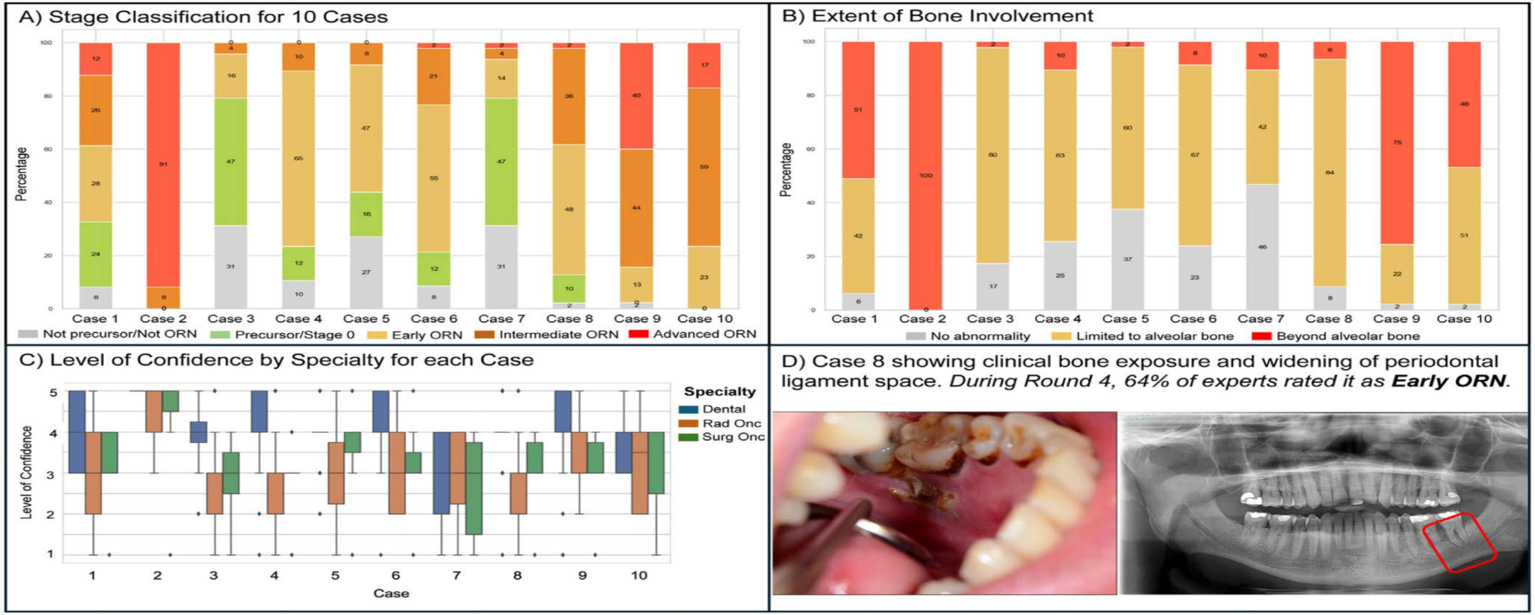
The Time Feature Conundrum



Should a time duration of exposed bone be required for diagnosis? The Consortium concluded: No (CS 2).

Moreno AC et al. IJROBP 2025 — Figure 2

Case-Based Classifications & Specialty Confidence



Panel A: Stage distribution. Panel B: Bone involvement. Panel C: Confidence by specialty. Panel D: Case 8.

Moreno AC et al. IJROBP 2025 — Figure 4

The Core Problem for AI/ML

Inconsistent outcome labeling makes AI/ML models unreliable.

If 69 international experts cannot consistently classify the same case, how can a machine learning model learn from retrospective data labeled by individual clinicians?

Standardization before prediction.

II

The Solution:

The ORAL Consortium &

Minimum Data Elements

The ORAL Consortium

Orodental Radiotherapy-Associated Late-Effects Consortium

69	4	93%	64%
International Experts	Delphi Rounds	Round 1 Participation	All 4 Rounds

Multidisciplinary: Radiation Oncology (n=26), Head & Neck Surgery/OMFS (n=12), Oral Medicine/Oncology & Dentistry (n=16), Medical Oncology, Radiology, Physics

Profile: 43% women | Mean age 47y | Median 16y in practice | 91% urban academic centers | Estimated 7% annual ORN incidence

Moreno AC et al. IJROBP 2025; 122(2):341–354

Delphi Consensus Process

Expert Email Invitation (n=75)

Excluded: Declined (n=2), No response (n=4)

↓

ORAL Consortium (n=69)

↓

Round 1: Review of existing ORN definitions and staging systems (n=64)

↓

Round 2: ORN terminology review, preliminary case & feature classifications (n=60)

↓

Round 3: ORN definition ranking, case & feature classifications (n=56)

↓

Round 4: Final ORN definition, repeat case & feature classification (n=54)

↓

Offline Panel Discussions during Manuscript Preparation

↓

The ORAL Consortium Consensus Report on ORN

Consensus-Based Definition of ORNJ

86% consensus

"A condition in which there is a loss of blood flow to bone tissue, which causes the bone to die. Findings of bone death may be clinical (exposed bone) and/or radiographic (sclerosis, pathologic fracture). It is caused by exposure to ionizing radiation and may occur at some point in time after radiation and in the absence of active disease in the site of bone death."

CS 1 Definition aligned with SNOMED-CT: bone disorder + radiation injury + vascular insufficiency + bone death

CS 2 No mandatory time duration of exposed bone for diagnosis

CS 3 Not all exposed bone equals necrotic bone

CS 4 ORNJ can be diagnosed with intact mucosa if radiographic evidence supports it

Key Consensus Statements

CS 5

CTCAE should be used in parallel with, not replace, an ORNJ classification system

CS 6

A time feature is not necessary for classifying severity; however, reporting time may complement monitoring duration or response to therapy

CS 7

Symptoms should not be disease state-defining; use validated PRO tools longitudinally

CS 8

All cases with pathologic fracture or fistula involving previously irradiated bone should be reported as Advanced ORNJ

CS 9

Adopt ORNJ-focused MDEs in multidisciplinary practice and clinical trial design

CS 10

Serial photographs + caliper measurements of exposed bone (mm) are strongly recommended

Moreno AC et al. IJROBP 2025; 122(2):341–354

The 9 Minimum Data Elements (MDEs)

Domain	Data Element	Example Value	SCTID	Type
Time	date_of_assessment	Date	439771001	Qualifier
Clinical	minor_bone_spicules	Present / Absent	367651003	Qualifier
Clinical	exposed_bone_length_mm	Numeric (mm)	30798006	Qualifier
Clinical	mucosal_status	Intact / Not intact	362224009	Qualifier
Clinical	probe_to_bone	Positive / Negative	386053000	Finder
Clinical	disorder_present	Fistula / Fracture / Ulcer	7907009	Finder
Radiographic	imaging_type (DICOM)	CT / OPG / MRI	363679005	Qualifier
Radiographic	morphology	Sclerosis / Osteolysis / Erosion	116676008	Morphology
Radiographic	vertical_ab_abnormality	Above / Below alveolar	3857009	Qualifier

Each MDE is coded with SNOMED-CT identifiers for AI/ML-ready data sets and health information exchange.

Moreno AC et al. IJROBP 2025 — Table 2

The ClinRad Classification System

Development and Standardization of an Osteoradionecrosis Classification System in Head and Neck Cancer: Implementation of a Risk-Based Model

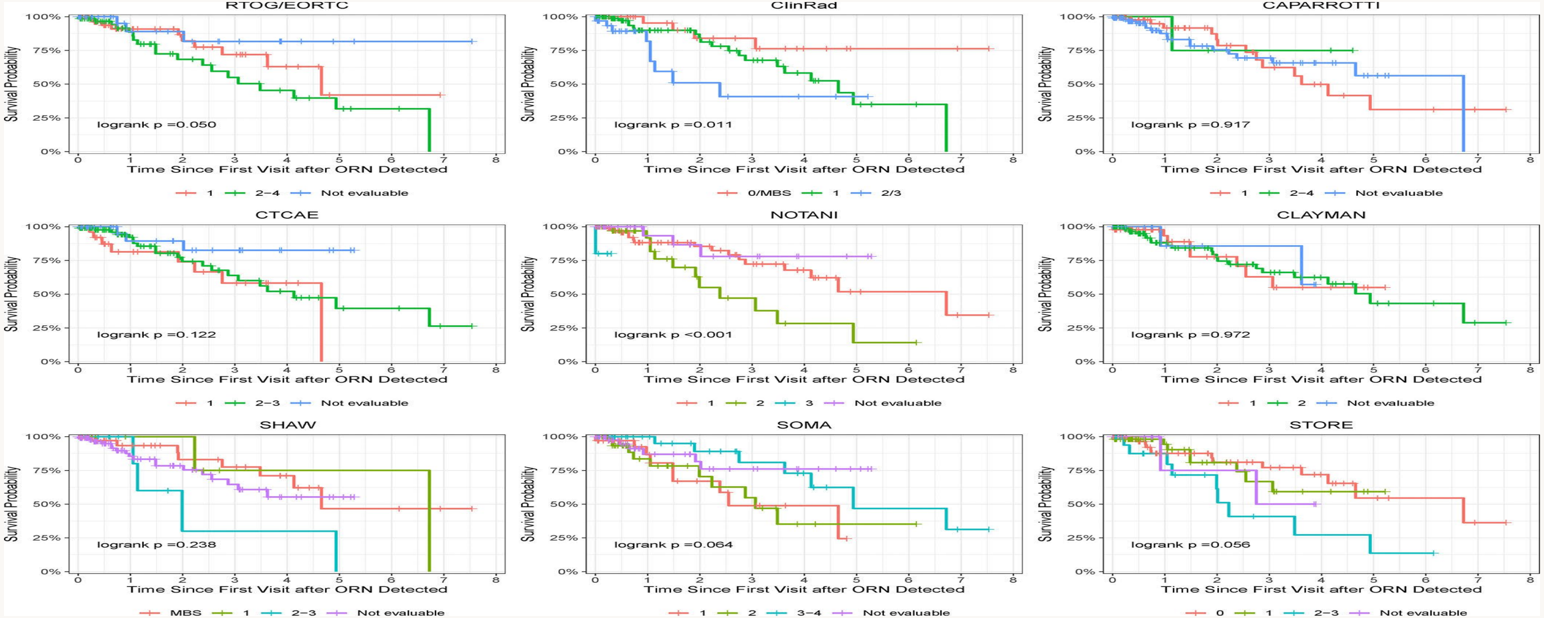
Erin E. Watson E, DMD, MSHc[1,2], Katrina Hueniken, MPH, MSc[3], Junhyung Lee, DDS[1], Shao Hui Huang, MD, MSc, MRT(T)[4], Amr El Maghrabi, DDS[1], Wei Xu, PhD[3], Amy C. Moreno, MD[5], C. Jillian Tsai, MD, PhD[4], Ezra Hahn, MD[4], Andrew J. McPartlin, MD[4], Christopher M.K.L. Yao, MD[6], David P. Goldstein, MD[6], John R. De Almeida, MD[6,7], John N. Waldon, MD[4], Clifton D. Fuller, MD[4], Andrew J. Hope, MD, PhD[4], Salvatore L. Ruggiero, DMD, MD[8,9], Michael Glogauer, DDS, PhD[2], Ali A. Hosni, MD, MSc, PhD[4]

Description	Stage	Radiographic Findings	Clinical Findings	Intervention	ClinRad clinical trial grade proposal
Distinct from ORN, occurs more often in patients with history of RT	**Minor Bone Spicules**	None, aside from superficial sequestra	Superficial mobile spicules/ sequestra within mucosa	None indicated	**Grade 0**
Radiographic evidence of bone necrosis confined to alveolar bone with no clinical signs of ORN	**Stage 0**	Bone necrosis confined to alveolar bone including: • Bone lysis/sclerosis • Widening periodontal ligament (PDL) space • Absence of osseous filling of extraction sockets	Intact Mucosa	None indicated	**Grade 1**
Clinical signs of ORN with or without radiographic evidence of bone necrosis confined to alveolar bone	**Stage 1**	None or as stage 0	Exposed bone[A]	Minor surgical intervention[B] and or medical management[C] may be indicated with or without adjunctive conservative management[D]	**Grade 2**
Radiographic evidence involving basilar bone with or without clinical signs of ORN	**Stage 2**	Bone necrosis involving the basilar bone or maxillary sinus	Intact mucosa or Exposed bone[A]	Intermediate surgical intervention may be indicated[E], with or without adjunctive conservative and medical management	**Grade 3**
Advanced ORN	**Stage 3**	One or more of the following, ○ Pathological fracture ○ Orocutaneous fistula ○ Oral antral communication/Oral nasal communication	One or more of the following, ○ Pathological fracture ○ Orocutaneous fistula ○ Oral antral communication/ Oral nasal communication	Reconstructive surgical intervention is indicated[F], with or without adjunctive conservative and medical management	**Grade 4**

0% missingness | Highest concordance (0.73) | Adopted by ISOO-MASCC-ASCO | 92% ORAL Consortium agreement

Watson EE et al. JCO 2024; 42(16):1922–1933

ClinRad Separability: Time-to-Serious ORN Event



KM curves for time to jaw fracture or surgery. ClinRad (p=0.011) outperforms most existing systems.

Watson EE et al. JCO 2024 — Figure 1

Missingness Comparison Across Systems

System	Total Visits	Non-Evaluable	% Missing	% Most Severe	Concordance
ClinRad	1063	0	**0%**	5.7%	**0.73**
CTCAE	1063	0	**0%**	10.4%	0.59
Notani	1063	34	3.2%	37.1%	0.68
Clayman	1063	445	41.9%	84.0%	0.50
Epstein	1063	429	40.4%	21.8%	0.59
Marx	1063	984	92.6%	67.1%	0.57
Coffin	1063	857	80.6%	7.3%	0.57
Glanzmann	1063	678	63.8%	5.8%	0.63

ClinRad: 0% missingness and 5.7% most-severe classification. Many existing systems have >40% missingness or overclassify severity.

Watson EE et al. JCO 2024 — Table 3

III

The Evidence:
AI/ML for ORN —
What Works and What Breaks

Dose-Volume Correlates of ORN

Study Design

Retrospective case-matched comparison

68 ORN cases, 131 matched controls

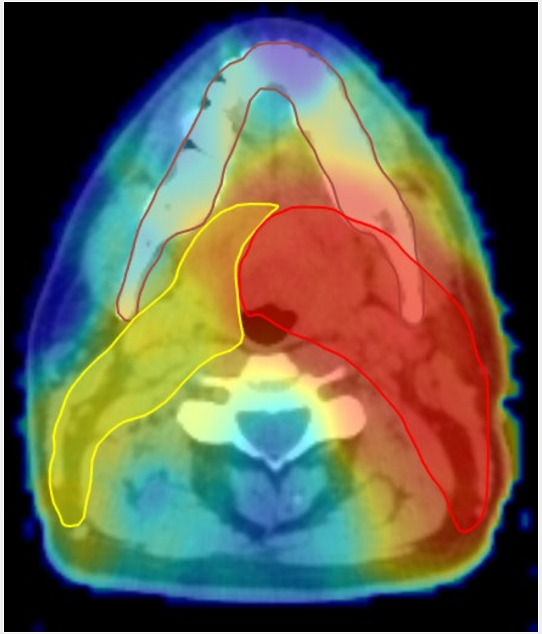
Oropharyngeal cancer patients, IMRT era

Mandibular DVH extraction per patient

Recursive Partitioning Analysis (RPA)

Key Question

With IMRT, a larger part of the mandible is subjected to beam-path doses unobserved in prior RT eras. What dose-volume parameters best correlate with ORN?



Mohamed ASR et al. Radiother Oncol 2017; 124:232–239

Dose-Volume Correlates of ORN: Findings

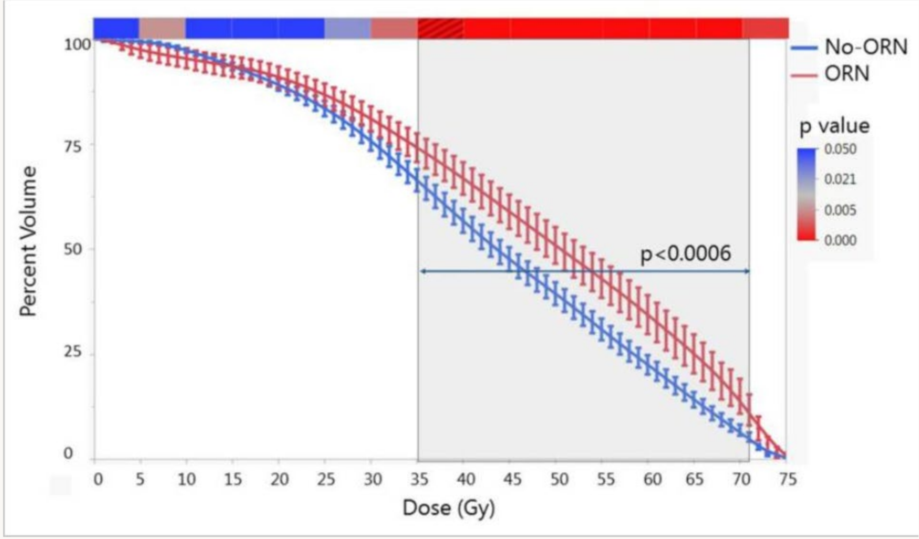
Key Results

Mean mandibular dose: 48.1 vs 43.6 Gy (p<0.0001)

Maximum dose: NOT significantly different

DVH bins V35–V73 all significant

RPA threshold: V44≥42% AND V58≥25%



Clinical Implication

Volume of bone receiving intermediate doses matters more than the maximum point dose. This shifted the paradigm from max dose constraints to

Limitation: single institution, consistent case ascertainment

Mohamed ASR et al. Radiother Oncol 2017; 124:232–239

NTCP Prediction Model for ORN

Normal Tissue Complication Probability (NTCP) Prediction Model for Osteoradionecrosis of the Mandible in Patients With Head and Neck Cancer After Radiation Therapy: Large-Scale Observational Cohort

Lisanne V. van Dijk, PhD,*,† Abdelrahman A. Abusaif, MD,*
Jillian Rigert, DMD, MD,* Mohamed A. Naser, PhD,*
Katherine A. Hutcheson, MD, PhD,*,‡ Stephen Y. Lai, MD, PhD,*,‡
Clifton D. Fuller, MD, PhD,* and Abdallah S.R. Mohamed, MD*
on behalf on the MD Anderson Symptom Working Group

First NTCP model for mandibular ORN

Study Design

- 1,259 HNC patients (2005–2015)
- 173 developed ORN (13.7% incidence)
- Bootstrapped forward variable selection
- 80/20 train-validation split

Variables Considered

DVH parameters.

Clinical factors:

- Dental extractions
- Smoking
- Chemotherapy
- Tumor site

van Dijk LV, Mohamed ASR et al. IJROBP 2021; 111(2):549–558

NTCP Prediction Model for ORN: Findings

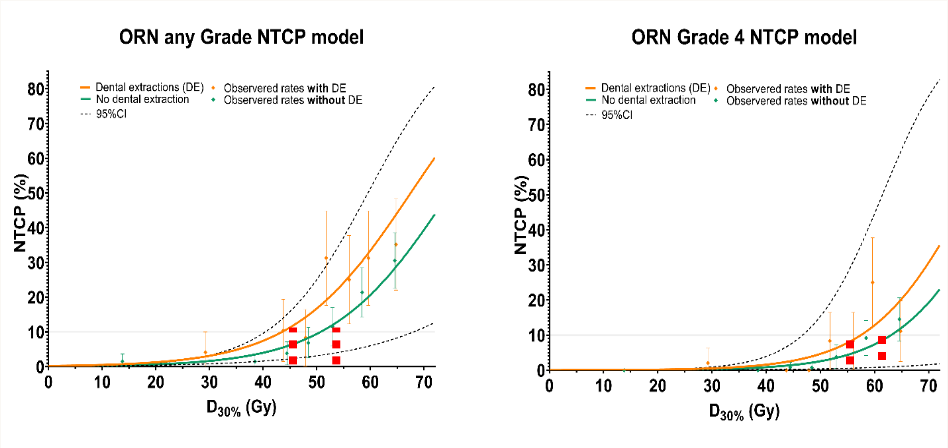
Key Results

Best predictors: Mandible D30% + pre-RT dental extraction

AUC (train): 0.78 (0.74–0.82)

AUC (validation): 0.75 (0.69–0.82)

Grade IV model AUC: 0.81 (train) | 0.82 (validation)



Clinical Threshold

<30% of mandible should receive ≥35 Gy for <5% ORN risk

Limitation: Single institution | ORN graded by institution-specific criteria

van Dijk LV, Mohamed ASR et al. IJROBP 2021; 111(2):549–558

PREDMORN: Multi-Institutional NTCP

Normal Tissue Complication Probability (NTCP) Prediction Model for Osteoradionecrosis of the Mandible in Patients With Head and Neck Cancer After Radiation Therapy: Large-Scale Observational Cohort

Lisanne V. van Dijk, PhD,*,† Abdelrahman A. Abusaif, MD,*
Jillian Rigert, DMD, MD,* Mohamed A. Naser, PhD,*
Katherine A. Hutcheson, MD, PhD,*,‡ Stephen Y. Lai, MD, PhD,*,‡
Clifton D. Fuller, MD, PhD,* and Abdallah S.R. Mohamed, MD*
on behalf on the MD Anderson Symptom Working Group

Largest multi-institutional ORN cohort worldwide

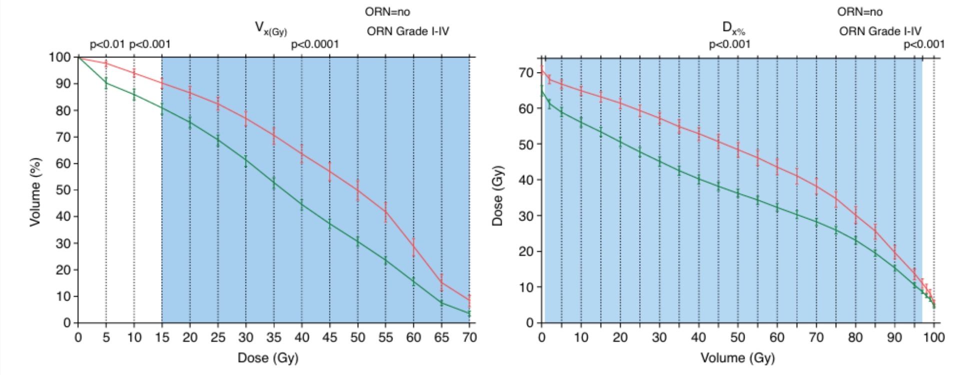
3,928 patients | 622 ORN cases | 8 institutions

Europe & North America (UK, US, Denmark, Netherlands, Spain, Canada)

Forward stepwise logistic regression with correlation-based preselection

80/20 stratified split + 2 external validation cohorts

External cohort 2: population-based (n=2,687)



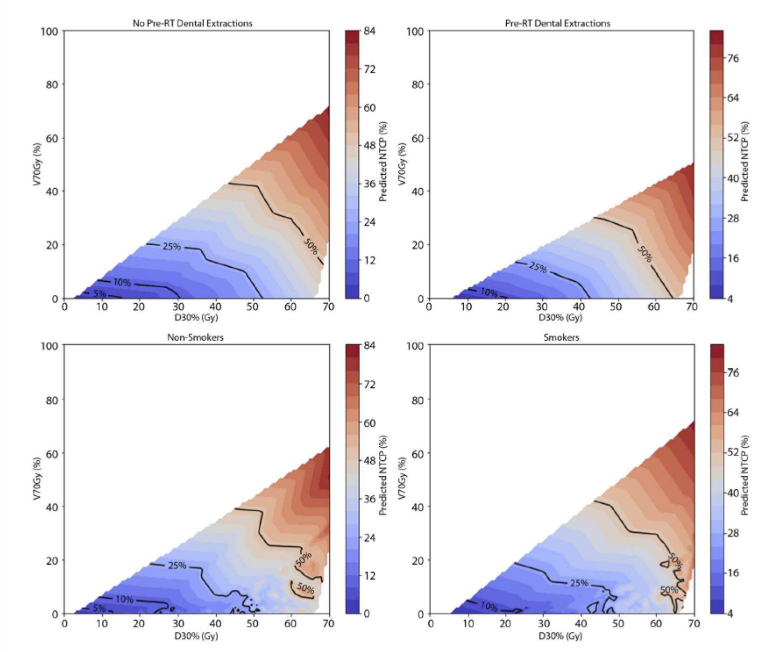
Key Challenge

Binary ORN endpoint due to inconsistent grading across sites (Notani vs Tsai vs CTCAE) — could not harmonize stages retrospectively

Humbert-Vidan L et al. IJROBP 2026 — PREDMORN Consortium

PREDMORN: Multi-Institutional NTCP — Findings

Variable	OR (95% CI)	P
D30% (Gy)	1.05 (1.03–1.07)	<.001
Pre-RT extraction	1.62 (1.21–2.15)	.001
V70Gy (%)	1.03 (1.00–1.05)	.024
Smoking (current)	1.39 (1.03–1.87)	.026



Validation:

AUC: 0.67 (train) | 0.69 (test) | 0.69 (ext1) | 0.65 (ext2)

Calibration: Brier 0.077, Log Loss 0.281 on pop-based cohort

OPC/Adv Larynx subcohort: AUC improved to 0.75

D30%≥45 Gy + V70Gy≥10% → 3-fold ORN risk increase (50.6% vs 16.8%)

Humbert-Vidan L et al. IJROBP 2026

Time-to-ORN Modeling

Externally validated digital decision support tool for time-to-osteoradionecrosis risk-stratification using right-censored multi-institutional observational cohorts

Laia Humbert-Vidan [m,1], Serageldin Kamel [m,1], Andrew Wentzel [b], Zaphanlene Kaffey [m], Moamen Abdelaal [c], Kyle B. Spier [d], Natalie A. West [m], G.Elisabeta Marai [b], Guadalupe Canahuate [e], Xinhua Zhang [b], Melissa M. Chen [m], Kareem A. Wahid [m], Jillian Rigert [m], Seyedmohammadhossein Hosseinian [f], Andrew J. Schaefer [g], Kristy K. Brock [m], Mark Chambers [m], Adegbenga O. Otun [m], Ruth Aponte-Wesson [m], Vinod Patel [h], Andrew Hope [i,j], Jack Phan [m], Adam S. Garden [m], Steven J. Frank [m], William H. Morrison [m], Michael T. Spiotto [m], David Rosenthal [m], Anna Lee [m], Renjie He [m], Mohamed A. Naser [m], Erin Watson [j], Katherine A. Hutcheson [m], Abdallah S.R. Mohamed [m], Vlad C. Sandulache [k], Lisanne V. van Dijk [l], Amy C. Moreno [m], Teresa Guerrero Urbano [h,*,2], Clifton D. Fuller [m,*,2], Stephen Y. Lai [m,*,2], For the MD Anderson Head and Neck Cancer Symptom Working Group

First model linking radiation dose to time-to-ORN onset

Weibull Accelerated Failure Time (AFT) model

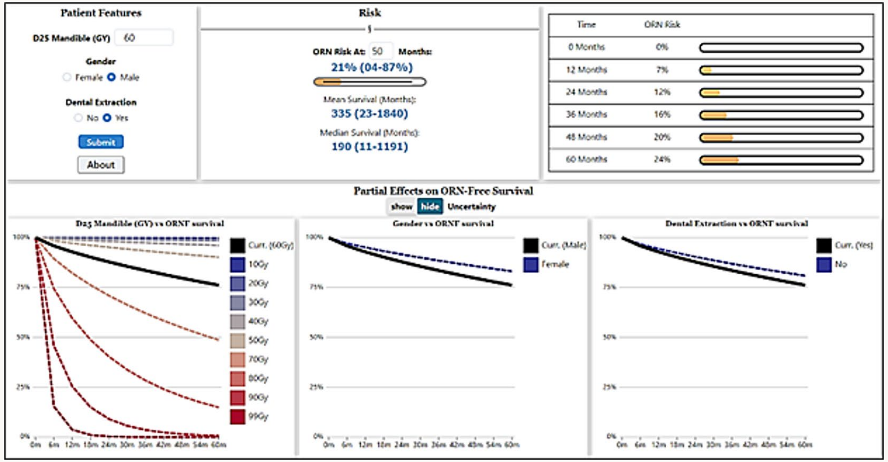
Training: 1,129 patients (198 ORN) at MD Anderson

External validation: 265 patients (92 ORN) at Guy's & St Thomas'

Covariates from updated NTCP: D25%, gender, pre-RT extraction

Innovation

Right-censored time-to-event analysis — accounts for patients still under surveillance who haven't developed ORN yet



Humbert-Vidan L et al. Radiother Oncol 2025; 207:110890

Time-to-ORN Prediction: Findings

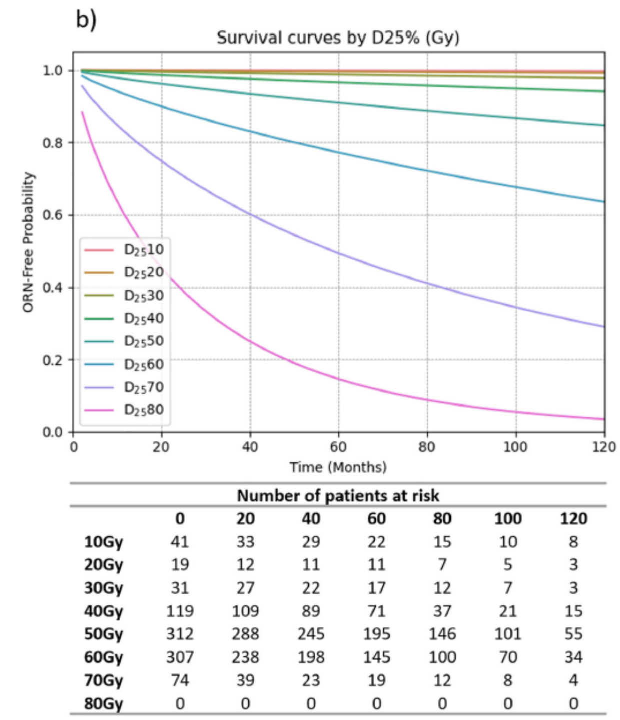
Covariate	ATR	Meaning	P
D25% (per Gy)	0.88	12% faster onset/Gy	<0.005
Gender (male)	0.62	38% faster onset	0.11
Pre-RT extraction	0.73	27% faster onset	0.13

Performance: C-index 0.72 at 72 months | IBS 0.133 | D-cal p=0.998

Shape ρ≈0.81: hazard decreases over time (highest risk early post-RT)

Online GUI: SUS usability score 85/100 across 25 users

Dose predicts not just IF but WHEN → risk-stratified surveillance



	0	20	40	60	80	100	120
10Gy	41	33	29	22	15	10	8
20Gy	19	12	11	11	7	5	3
30Gy	31	27	22	17	12	7	3
40Gy	119	109	89	71	37	21	15
50Gy	312	288	245	195	146	101	55
60Gy	307	238	198	145	100	70	34
70Gy	74	39	23	19	12	8	4
80Gy	0	0	0	0	0	0	0

Humbert-Vidan L et al. Radiother Oncol 2025; 207:110890

ML vs Deep Learning for ORN

Hypothesis

Can DL using full 3D dose distributions outperform traditional ML using dose summary statistics?

Study Design

ML arm: n=1,259 — LR, RF, SVM + random classifier

DL arm: n=1,236 — 3D ResNet, DenseNet, autoencoder on mandible dose maps

173 ORN+ / 1,086 ORN−

Same institution, same outcome labels

Comparison of Machine-Learning and Deep-Learning Methods for the Prediction of Osteoradionecrosis Resulting From Head and Neck Cancer Radiation Therapy

Brandon Reber, BS,[a,*] Lisanne Van Dijk, PhD,[a,b] Brian Anderson, PhD,[a,c] Abdallah Sherif Radwan Mohamed, MD, PhD,[a] Clifton Fuller, MD, PhD,[a] Stephen Lai, MD, PhD,[a] and Kristy Brock, PhD[a]

Table 2 Mean (±SD) metric values for the cross-validation withheld folds for the ML models*

Model	Accuracy	Balanced accuracy	Recall	Precision	F1 score	AUROC	AUPRC
Logistic regression	0.69 ± 0.05	0.70 ± 0.07	0.72 ± 0.14	0.27 ± 0.05	0.39 ± 0.07	0.74 ± 0.07	0.28 ± 0.08
Random forest	0.65 ± 0.05	0.69 ± 0.07	0.74 ± 0.14	0.25 ± 0.04	0.37 ± 0.06	0.69 ± 0.07	0.23 ± 0.04
Support vector machine	0.69 ± 0.04	0.70 ± 0.07	0.71 ± 0.13	0.27 ± 0.04	0.39 ± 0.06	0.70 ± 0.07	0.24 ± 0.04
Random classifier	0.52 ± 0.04	0.49 ± 0.08	0.45 ± 0.14	0.14 ± 0.04	0.21 ± 0.07	0.50 ± 0.00	0.14 ± 0.01

Abbreviations: AUPRC = area under the precision recall curve; AUROC = area under the receiver operating characteristic curve; ML = machine learning.

* Each cell shows the mean (±SD) of the metrics from the withheld folds of the stratified 10-fold cross-validation with 10 repeats.

Reber B, van Dijk LV, Mohamed ASR et al. Adv Radiat Oncol 2023

ML vs Deep Learning for ORN

DL did NOT outperform traditional ML.

Interpretation

The authors conclude DL underperformance reflects:

- Insufficient training data for DL (no improvement trend 10%–100%)
- Low event rate (13.7%) exacerbates class imbalance for DL
- Dose map low-level features may not be as powerful as DVH associations

Table 3 **Performance of the best DL models for each architecture type***

Architecture	Accuracy	Balanced accuracy	Recall	Precision	F1 score	AUROC	AUPRC
ResNet	0.87	0.69	0.04	0.50	0.07	0.57	0.23
DenseNet	0.83	0.54	0.10	0.21	0.14	0.58	0.17
Autoencoder	0.71	0.53	0.33	0.18	0.23	0.59	0.15
Random	0.49	0.46	0.46	0.11	0.17	0.49	0.13

Abbreviations: AUPRC = area under the precision recall curve; AUROC = area under the receiver operating characteristic curve; DL = deep learning.
* The reported metrics are from the withheld test set not used during model training or selection. Metrics sensitive to data imbalance, such balanced accuracy, F1 score, and AUPRC, were lower than those for the logistic regression model using the test set.

Scaling DL for ORN requires both larger datasets and, as we will see, standardized outcome labels to enable multi-institutional training.

Reber B et al. Adv Radiat Oncol 2023; 8(4):101163

Unsupervised ML: Whole-DVH Clustering

Cluster-Based Toxicity Estimation of Osteoradionecrosis Via Unsupervised Machine Learning: Moving Beyond Single Dose-Parameter Normal Tissue Complication Probability by Using Whole Dose-Volume Histograms for Cohort Risk Stratification

Seyedmohammadhossein Hosseinian, PhD,* Mehdi Hemmati, PhD,† Cem Dede, MD, MSc,‡ Travis C. Salzillo, PhD,‡ Lisanne V. van Dijk, PhD,§ Abdallah S.R. Mohamed, MD, PhD,‡,‖ Stephen Y. Lai, MD, PhD,¶ Andrew J. Schaefer, PhD,# and Clifton D. Fuller, MD, PhD‡,#

Study Design

K-means clustering on whole mandibular DVH shapes

n = 1,259 patients (173 ORN, 13.7%)

Soft-margin SVM to partition dose-volume space

Risk indices per region based on ORN incidence and PDE status

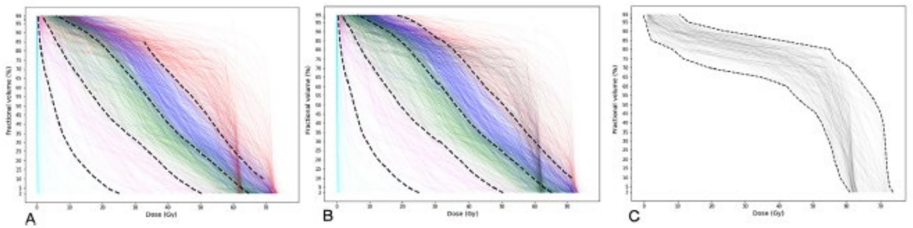


Fig. 3. Cluster borders identified by soft-margin support vector machine. (A) K = 5: Cluster borders; (B) K = 6: Borders of the first 5 clusters; (C) K = 6: Envelopes of the sixth cluster.

Innovation

Evaluates ORN risk based on whole DVH shape rather than single-parameter summary (e.g., D30%); addresses multicollinearity while preserving clinical interpretability

Hosseinian S, Mohamed ASR et al. IJROBP 2024; 119(5):1569–1578

Unsupervised ML: Whole-DVH Clustering: Findings

6 DVH clusters identified with graded risk indices from 0% to ~30% per PDE status

Key Findings

DVH shape matters: patients with same D30% (55–60 Gy) had different ORN rates by cluster (13.6% vs 23.5%)

Cross-validation: actual ORN fell within 95% CI of training estimates for all but one cluster-PDE bin

Provides visual risk-assessment tool and dose constraints for optimization; however, DVH-based models remain limited by loss of spatial information and class inseparability reflecting the multifactorial pathogenesis of ORN.

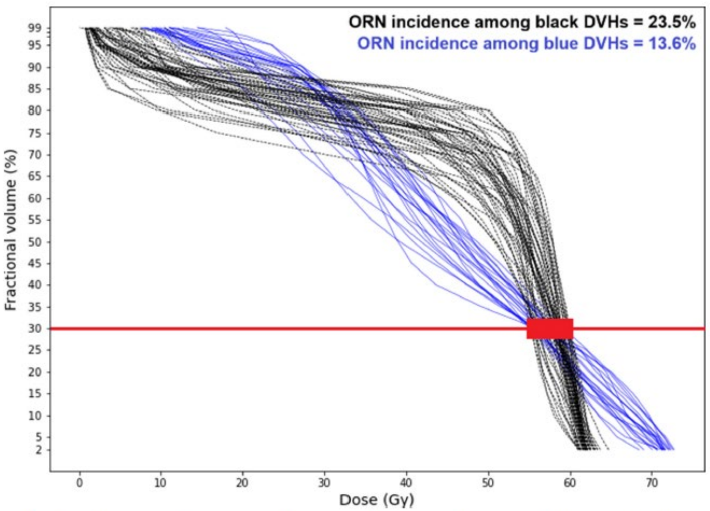


Fig. 5. Different osteoradionecrosis incidences among dose-volume histograms with the same D30% value.

Hosseinian S et al. IJROBP 2024; 119(5):1569–1578

The Recurring Bottleneck

Across all modeling studies, the same limitations recur:

Heterogeneous Outcomes

Different institutions use different ORN grading → label noise

Limited Generalizability

Single-institution models fail external validation

DL ≠ Better Than ML

Richer features can't overcome inconsistent labels

Dose Alone Is Insufficient

ORN/non-ORN patients highly inseparable by dose

Model performance hits a ceiling without standardized labels.

CT Radiomics for ORN Detection

Computed tomography radiomics-based cross-sectional detection of mandibular osteoradionecrosis in head and neck cancer survivors

Serageldin Kamel[a], Laia Humbert-Vidan[a], Zaphanlene Kaffey[a], Sarah Mirbahaeddin[a], Abdulrahman Abusaif[a], David T.A. Fuentes[b], Kareem Wahid[a,b], Cem Dede[a], Mohamed A. Naser[a], Renjie He[a], Ahmed W. Moawad[c], Khaled M. Elsayes[c], Melissa M. Chen[c], Adegbenga O. Otun[d], Jillian Rigert[a], Mark S. Chambers[d], Andrew Hope[e], Erin Watson[e,f], Kristy K. Brock[b], Katherine Hutcheson[d], Lisanne van Dijk[g], Amy C. Moreno[a], Stephen Y. Lai[a,d,*], Clifton D. Fuller[a,*], Abdallah S.R. Mohamed[a,h,*], MD Anderson Head and Neck Cancer Symptom Working Group

Study Design

Cross-sectional radiomics-based ORN detection

150 patients with confirmed ORN (2008–2018, MDACC)

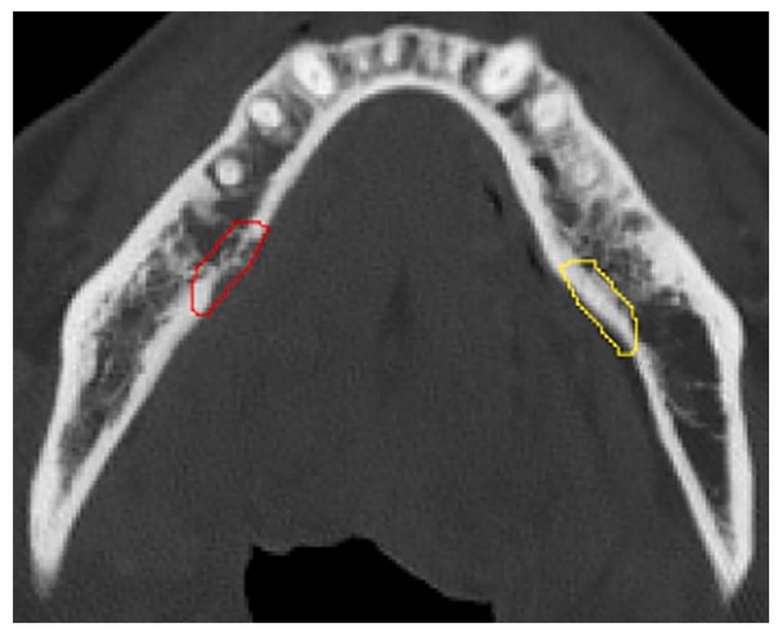
CECT images with manually segmented ORN regions

Contralateral healthy mandible as internal control

Pipeline

PyRadiomics extraction → correlation filtering → RF classifier with RFE

1,316 features → 506 after collinearity → 67 optimal



Kamel S, Mohamed ASR et al. Oral Oncol 2025; 167:107337

CT Radiomics for ORN Detection: Findings

Accuracy: 88% | AUC: 0.96

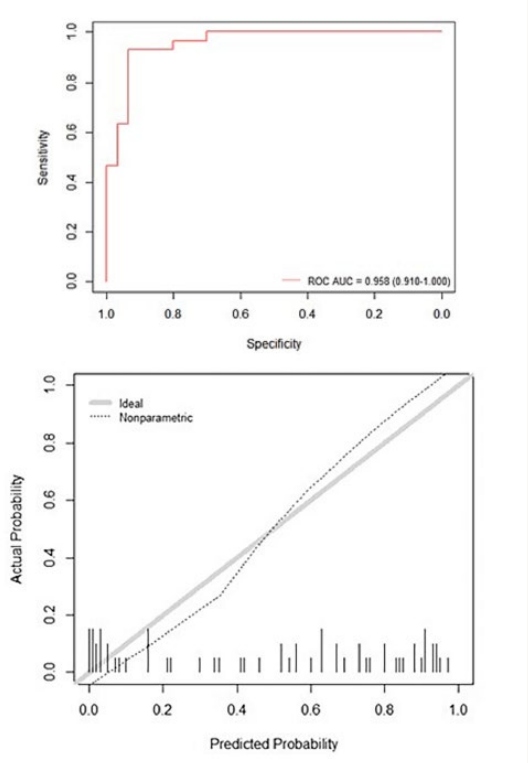
SHAP Explainability

Wavelet-LLH First-order Mean/Median most associated with ORN regions

Future Direction

Moving from detection → prediction: identifying subclinical ORN regions before clinical manifestation

Limitation: trained on confirmed ORN using institutional criteria



Kamel S et al. Oral Oncol 2025; 167:107337

DCE-MRI Quantitative Parameters for ORN

Quantitative Dynamic Contrast-Enhanced MRI Identifies Radiation-Induced Vascular Damage in Patients With Advanced Osteoradionecrosis: Results of a Prospective Study

Joint Head and Neck Radiation Therapy-MRI Development Cooperative; Abdallah S.R. Mohamed, MD, MSc,* Renjie He, PhD,* Yao Ding, PhD,* Jihong Wang, PhD,* Joly Fahim, BS,* Baher Elgohari, MD, MSc,* Hesham Elhalawani, MD, MSc,* Andrew D. Kim, BS,† Hoda Ahmed, MD, MSc,† Jose A. Garcia, MSN,† Jason M. Johnson, MD,‡ R. Jason Stafford, PhD,§ James A. Bankson, PhD,§ Mark S. Chambers, DMD, MS,† Vlad C. Sandulache, MD, PhD,‖ Clifton D. Fuller, MD, PhD,* and Stephen Y. Lai, MD, PhD*,†

Study Design

Prospective study: 30 patients with advanced ORN (Grade IV) requiring surgery

DCE-MRI with Ktrans and Ve parameter extraction

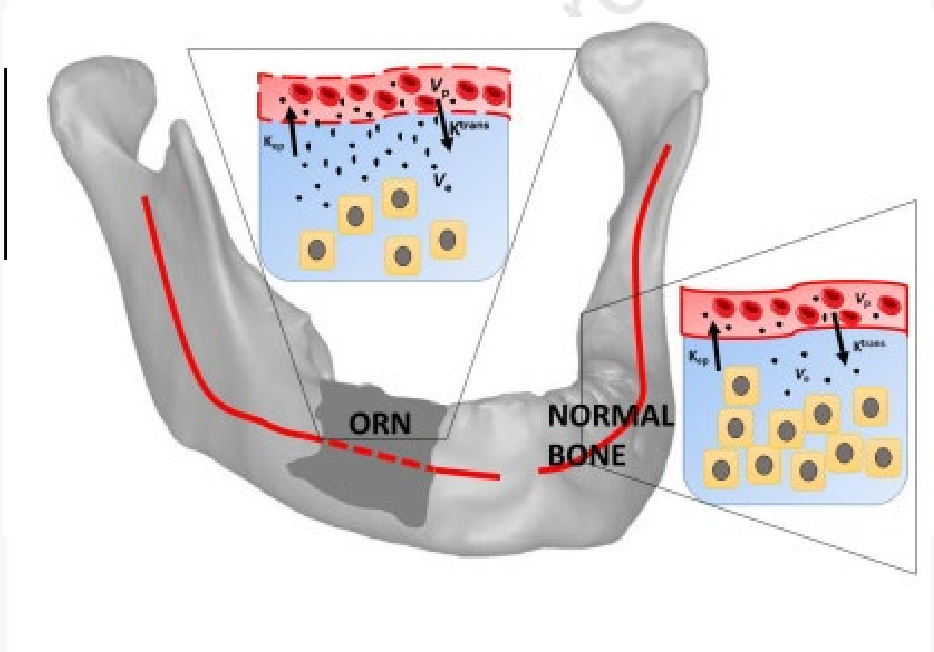
ORN vs contralateral normal mandible (internal control)

Median time to ORN diagnosis: 38 months post-IMRT

Key Question

Can quantitative DCE-MRI vascular parameters distinguish ORN from normal mandible?

Tests the radiation-induced hypocellular-hypovascular-hypoxic pathogenesis theory



DCE-MRI Quantitative Parameters: Findings

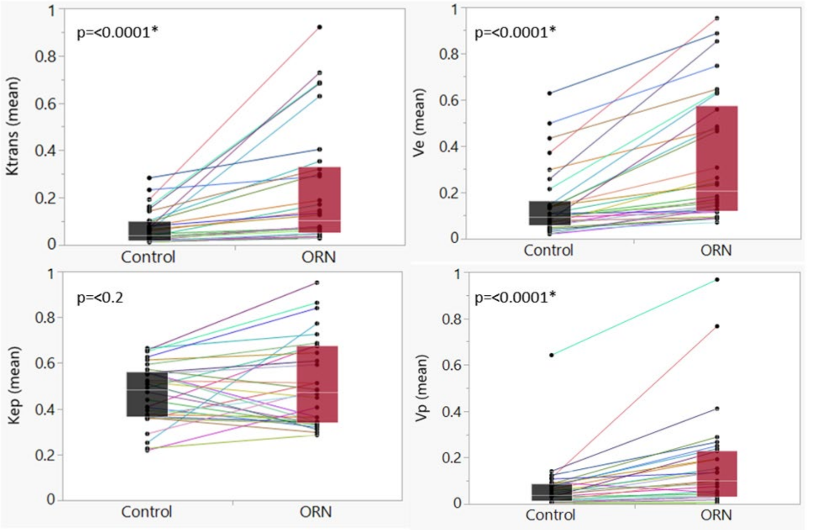
Ktrans 3.2x higher | Ve 2.7x higher in ORN vs control

Significance

Quantitatively confirms vascular leakiness in mandibular ORN regions, consistent with radiation-induced hypovascular pathogenesis

Future Direction

MRI-based vascular biomarkers for early subclinical ORN detection before clinical manifestation



Limitation: advanced ORN only (Grade IV); single institution

Mohamed ASR et al. IJROBP 2020; 107(3):552–558

IMPACT: Orodental Auto-Segmentation

Image-based mandibular and maxillary parcellation and annotation using computed tomography (IMPACT): a deep learning-based clinical tool for orodental dose estimation and osteoradionecrosis assessment

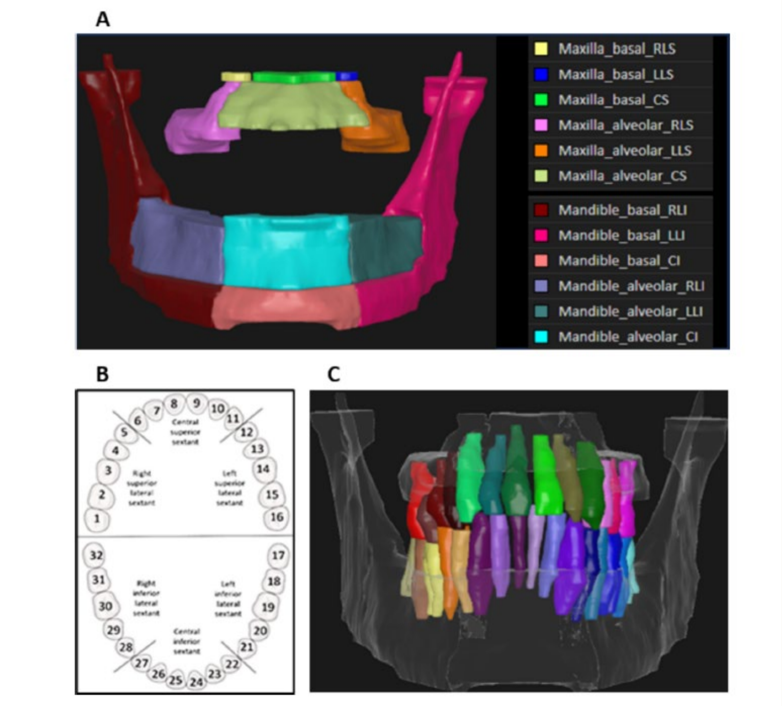
Laia Humbert-Vidan[a], Austin H. Castelo[b], Renjie He[a,1], Lisanne V. van Dijk[c], Dong Joo Rhee[d], Congjun Wang[d], He C. Wang[d], Kareem A. Wahid[b], Sonali Joshi[a], Parshan Gerafian[a], Natalie West[a], Zaphanlene Kaffey[a], Sarah Mirbahaeddin[a], Jaqueline Curiel[a], Samrina Acharya[a], Amal Shekha[a], Praise Oderinde[a], Alaa M.S. Ali[a], Andrew Hope[e], Erin Watson[f], Ruth Wesson-Aponte[g], Steven J. Frank[a,1], Carly E.A. Barbon[g,1], Kristy K. Brock[b,1], Mark S. Chambers[g,1], Muhammad Walji[h], Katherine A. Hutcheson[g,1], Stephen Y. Lai[g,1], Clifton D. Fuller[a,1], Mohamed A. Naser[a,1], Amy C. Moreno[a,*,1], for the OPC SURVIVOR Program, the MD Anderson Head and Neck Cancer Symptom Working Group

Study Design

DL auto-segmentation of orodental structures on planning CT

60 patients, 50 train / 10 test, 5-fold cross-validation

10 observers for manual ground truth contours



Humbert-Vidan L et al. Phys Imaging Radiat Oncol 2025; 35:100817

Three Segmentation Tasks

1. Mandible sub-volumes: Swin UNETR (one-stage)

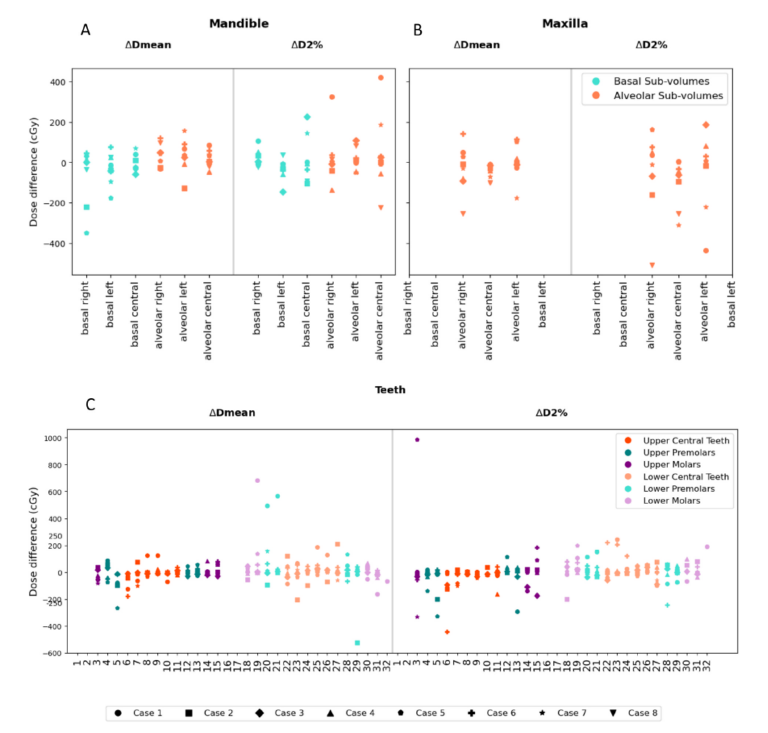
2. Maxilla sub-volumes: ResUNet → Swin UNETR (two-stage)

3. Individual teeth: ResUNet → Swin UNETR (two-stage)

ClinRad Alignment: alveolar vs basal distinction enables staging-concordant dosimetry

IMPACT: Orodental Auto-Segmentation: Findings

Region	Dice	HD95
Mandible basal	0.85	1.2
Mandible alveolar	0.82	1.5
Maxilla alveolar	0.78	1.6
Upper central teeth	0.80	1.4
Lower central teeth	0.76	1.6
Premolars	0.69–0.71	2.3
Molars	0.63–0.76	3.6



Dose-volume validation: most within ±2.5 Gy for Dmean

Enables spatially informed NTCP models and ClinRad-concordant image-based ORN detection

Humbert-Vidan L et al. PHIRO 2025; 35:100817

IV

The Path Forward:
MDEs as Infrastructure for
Next-Generation Models

MDEs as the Ground Truth Layer

Standardized MDEs ▶ Prospective Registries ▶ Multi-Institutional Data Sharing ▶ Federated Learning ▶ Validated Clinical DSS

What MDEs Enable:

- Consistent phenotyping across institutions for reliable training labels
- Scalable health information exchange via SNOMED-CT coded elements
- Human- and machine-readable ontology (ORNJ ontology under development)
- Stage migration tracking across evolving classification systems
- Open data enablement: standardized datasets for reproducible science

Broader Relevance: ORN & the ONJ Ecosystem

While this talk focuses on ORN, the principles are universal:

ClinRad was explicitly modeled after AAOMS staging for MRONJ to harmonize dental professional familiarity

Atomic MDEs are transferable: mucosal_status, probe-to-bone, morphology apply equally to MRONJ

The standardization-before-prediction principle applies to any disease with competing classification systems

Session context: this complements the MRONJ prediction research presented in this session

Summary: Key Messages

1. ORNJ has 9+ definitions and 16+ classification systems creating label noise for AI/ML
2. The ORAL Consortium (69 experts) established 10 consensus statements and 9 MDEs
3. ClinRad achieves 0% missingness and highest concordance motivating AI tools
4. A decade of ML/DL + PREDMORN (n=3,928) confirm performance ceiling without standardized outcomes
5. MDEs coded with SNOMED-CT are the foundation for multi-institutional, spatio-temporal predictive models

Standardization is not a barrier to innovation

—

it is the prerequisite.

Call to Action for the Dental/Oral/Craniofacial Community:

Adopt MDEs now
integrate into clinical workflows and trial design

Collect prospectively
serial assessments with photographs and quantitative measurements

Enable the models of tomorrow
consistent labels unlock multi-institutional AI/ML

Acknowledgments: The International ORAL Consortium | PREDMORN Consortium | NIDCR | NIH | Co-Investigators at MDA & BCM

abdallah.mohamed@bcm.edu | @AbdallahMohamed