Introducing Plan Quality Metrics

A Method to Customize and Automate the Measurement of Plan Quality

Ben Nelms, Ph.D. (Canis Lupus LLC)

Disclaimers

- **Sources of Funding**
  - None to report, i.e. the research and work discussed here are unfunded.
  - Plan Challenges are voluntary and free to participants.
  - Plan Challenge design, evaluation, and analysis is a voluntary service provided by me and the team of dosimetrists from ROR.
- **Statements of Conflict of Interest**
  - I am a paid consultant to Sun Nuclear Corporation and the inventor of their products: EPIDose, MotionSim, and 3DVH.
  - My company (Canis Lupus LLC) owns and develops the medical device software solution Quality Reports [EMR]®.

ACKNOWLEDGEMENTS

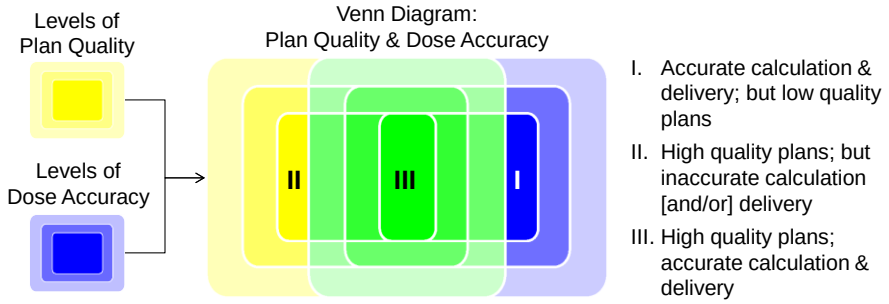
- **BELOW ARE SOME OF MY COLLEAGUES WHO WERE INSTRUMENTAL IN THIS WORK:**
  - Greg Robinson, MS, CMD, RT(T)
  - Kyle Velasco, CMD, RT(T)
  - Adam Moore, CMD, RT(T)
  - Steve Boyd, CMD, RT(R)(T)
- **OTHER CONTRIBUTORS INCLUDE:**
  - Vladimir Feygelman, Ph.D.
  - "Project Icarus" research team (consists of a small, international group of dosimetrists and physicists from the USA, CAN, AUS, and the UK)

OUTLINE

- **INTRODUCTION: WHAT IS PLAN QUALITY?**
- **PLAN QUALITY METRICS (PQM)**
  - History: The International "Plan Challenges"
  - Description & Methods
  - Example Plan Challenge Results / Analysis
- **APPLICATIONS (PER PATIENT)**
  - Per-Patient Workflow
  - Commissioning & Validation
  - Accreditation & Competency Testing
- **OTHER APPLICATIONS (QUALITY MANAGEMENT SYSTEMS)**
  - Commissioning (TPS and Delivery)
  - Benchmarking & Comparative Effectiveness
  - Clinical Trials

What is Plan Quality?

- **Plan Quality ≠ Dose Accuracy (calc & delivery)**
  - You may be able to accurately calculate and deliver a plan…but if it's a low quality plan to begin with, you produce a low quality result.
  - TPS dose algorithm accuracy ≠ plan quality
  - Machine delivery accuracy ≠ plan quality



What is Plan Quality?

- **Plan Quality ≠ Any Particular Modality, Brand, or Acronym**
  - **"IMRT"** does not guarantee high quality plans
  - **"VMAT"** does not guarantee high quality plans
  - **"Particle/Proton Therapy"** does not guarantee high quality plans
  - **[Insert New Fancy Product Name Here]** does not guarantee high quality plans
- **Findings from the Plan Challenges**
  - Plenty of poor quality plans using latest modalities and products
  - Some of the very high quality plans are some of the least complex and using older equipment

What is Plan Quality?

- **Plan Quality ≠ Planner Experience or Certification**
  - **Years Experience** does not guarantee high quality plans
  - **CMD** does not guarantee high quality plans (nor do: PhD, DABR, MD, etc.)
  - Currently, there is no objective testing of practical skills (i.e. contouring or planning) included in the CMD exams.
- **Findings from the Plan Challenges**
  - Plenty of poor quality plans from very experienced and certified planners.
  - Some of the highest quality plans have come from brand new (< 1 year) planners and dosimetry students.

What is Plan Quality?

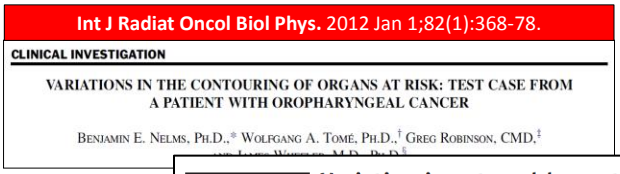
- **Definition of Plan Quality**

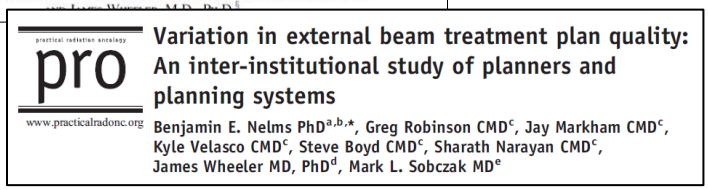
plan qual·i·ty ['plan 'kwä-lə-tē]

1. The objective measure of how well a 3-D dose distribution, when coupled with 3-D anatomy, meets clearly defined goals and priorities.

History: The "Plan Challenge"

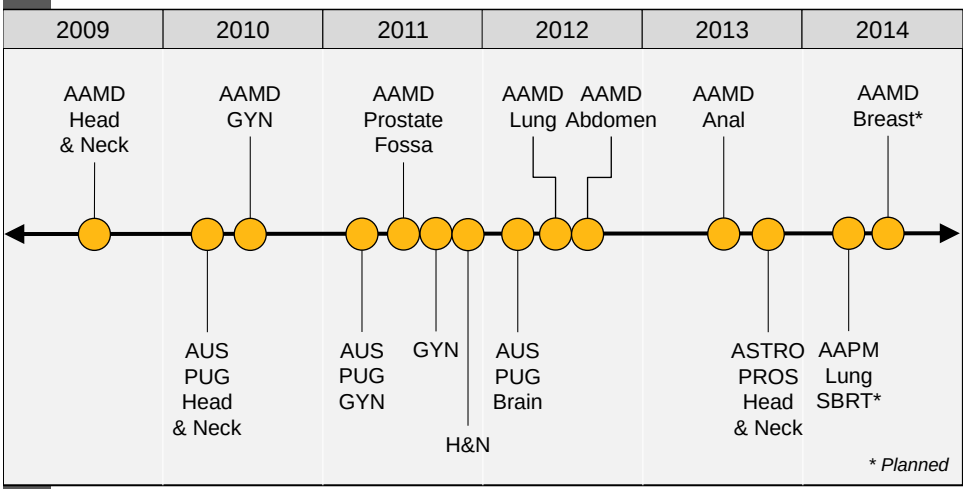
- **The Mission of the Plan Challenge Initiative**
  - To perform controlled, scientifically-valid studies of the variation in Plan Quality across treatment planners and modalities, with the aims to: glean best practices, educate our peers, improve quality in radiation therapy, and inspire continual improvement.





History: The "Plan Challenges"

- **Plan Challenge Timeline**

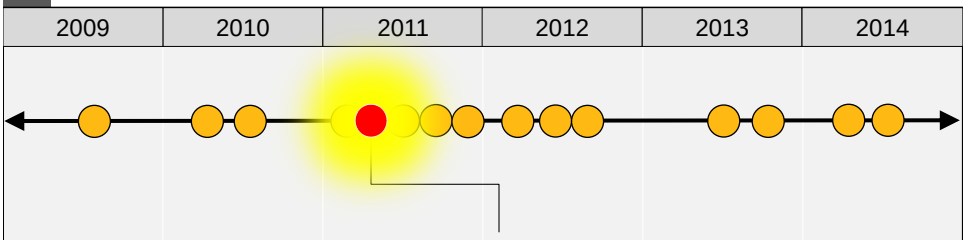


History: The "Plan Challenges"

- **Plan Challenge Accrued Database**
  - Over 10 different test datasets
    - CT imageset with required contours (provided)
    - Plan Quality Algorithms (i.e. Objectives & Scoring Methods)
  - Over 1800 submitted plans
    - DICOM RT Plan & Dose pairs
  - Over 30,000 metrics
    - Total Score (PQM) for each submitted dataset
    - Sub-metric results and sub-scores per metric
    - Performance distributions over population of planners

History: The "Plan Challenges"

- **"The Conversation"**



***The Conversation*:**
Can we measure Plan Quality in a way that is:
1) QUANTITATIVE, 2) COMPREHENSIVE, 3) CLEAR, 4) INARGUABLE, and 5) FULLY TRANSPARENT (i.e. fair)?

METHOD: "PLAN QUALITY ALGORITHM"

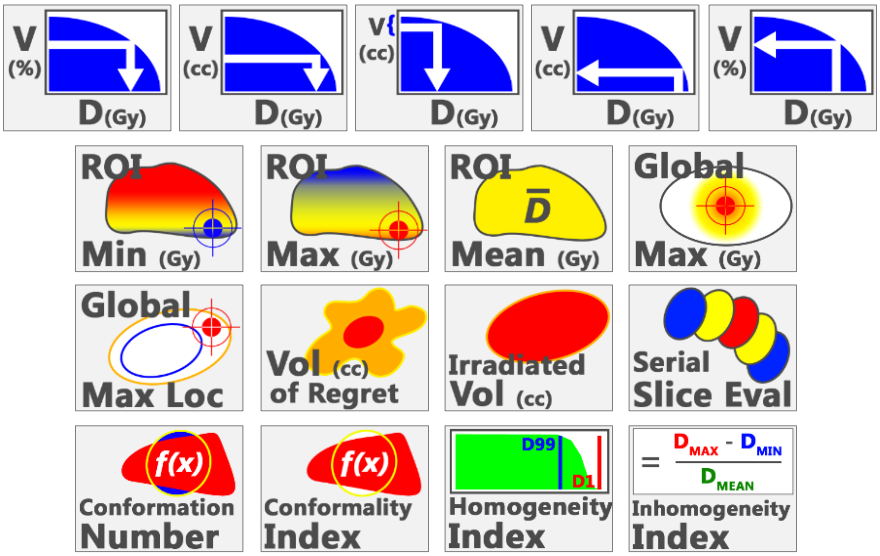
DEFINE PLAN QUALITY ALGORITHM

- **IDENTIFY CRITICAL SUB-METRICS.** Dose, DVH, or formulaic sub-metrics selected from a library of choices (currently 17 options)
- **DEFINE EACH SUB-METRIC'S PARAMETERS.** ROI, dose and/or volume levels, etc. Can also set "ROI Synonyms" to allow for some variability in ROI naming.
- **DEFINE EACH SUB-METRIC'S SCORE FUNCTION.** Specify priority (i.e. weight) along with "failure" level and a "goal" (e.g. ideal) level, and scoring in between.

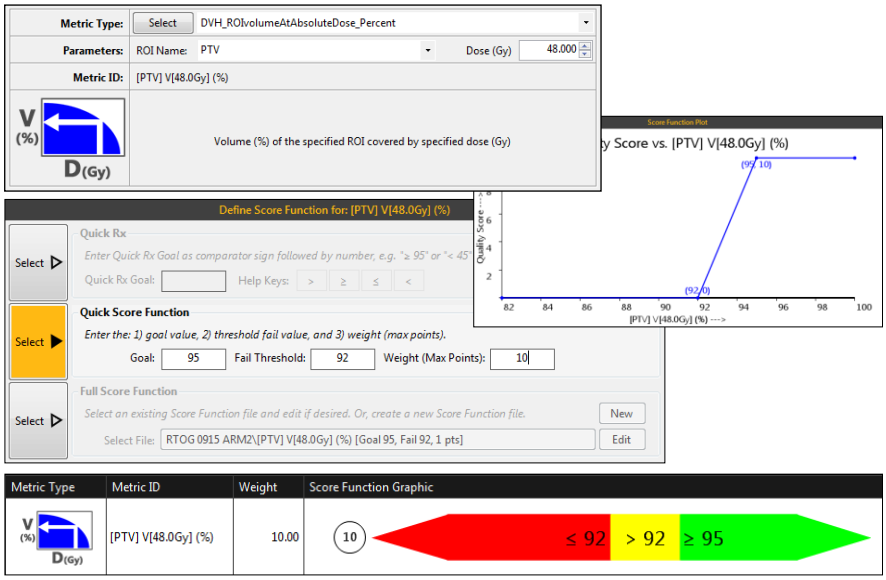
SCORE PLAN

- **IMPORT DICOM DATA.** RT Plan, Structures, Dose, and CT images.
- **LOAD PLAN QUALITY ALGORITHM.** Generates score automatically along with full spreadsheet and per-metric "drill down" analysis.

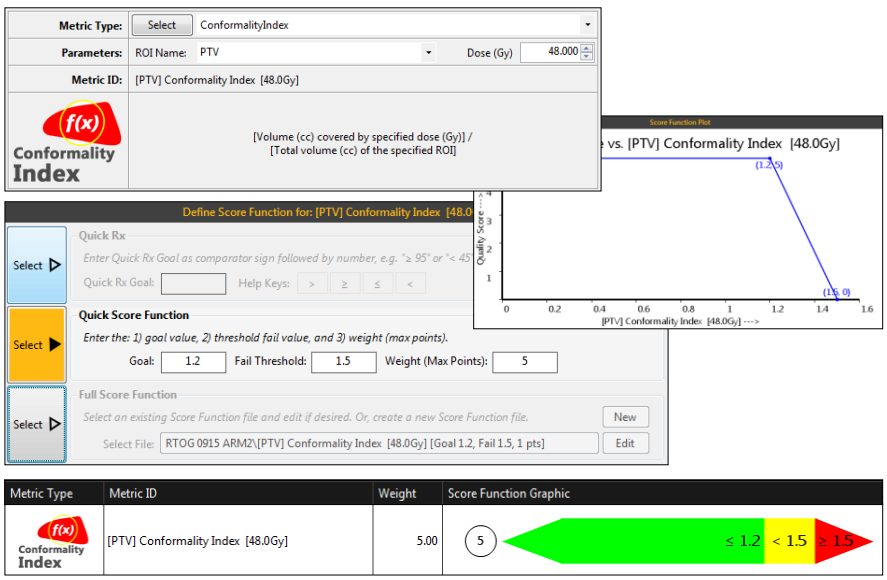
LIBRARY OF PLAN QUALITY SUB-METRICS



EXAMPLES: DEFINING SUB-METRICS

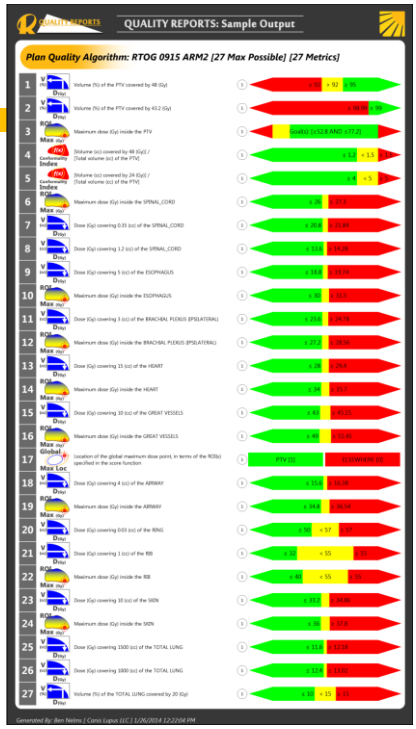


EXAMPLES: DEFINING SUB-METRICS



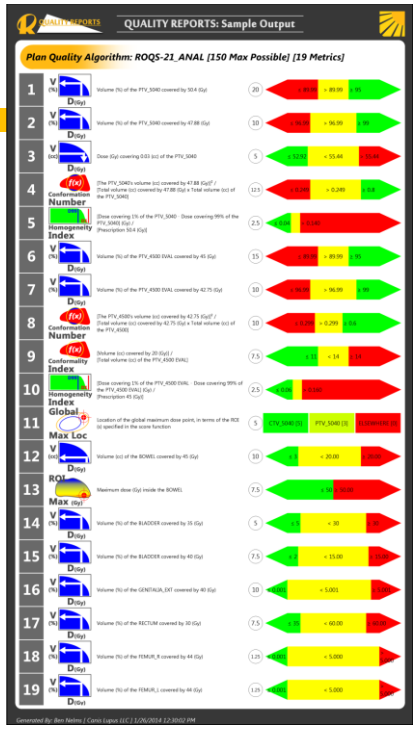
EXAMPLE: RTOG 0915

- **RTOG 0915 (ARM 2) RENDERED AS PQ ALGORITHM**
  - 27 sub-metrics, each weighted equally (1.00)
  - 26 sub-metrics have “ideal” levels along with “acceptable” levels; one is pass/fail
  - 15 DVH-based
  - 9 Simple (i.e. min, max mean, etc.)
  - 3 Advanced or Formulaic (i.e. conformality indices, global max location, etc.)



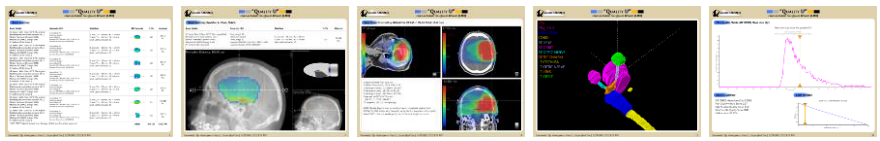
EXAMPLE: 2013 PC

- **2013 PLAN CHALLENGE PQ ALGORITHM**
  - 19 sub-metrics, variably weighted for a total score of 150.00
  - All 19 sub-metrics have “ideal” levels along with “acceptable” levels
  - 12 DVH-based
  - 1 Simple (i.e. min, max mean, etc.)
  - 6 Advanced or Formulaic (i.e. conformation numbers, conformality indices, homogeneity indices, global max location, etc.)

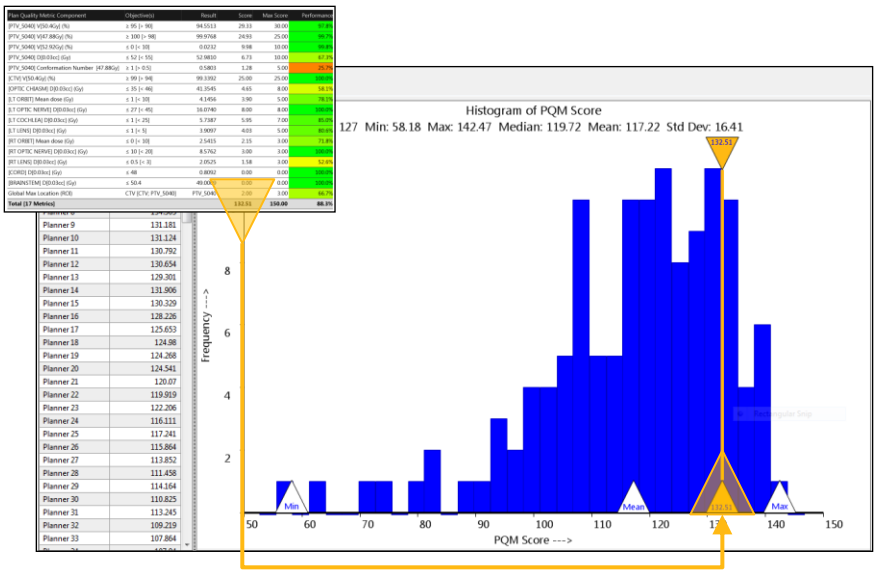


GENERATING PQM SCORE

Plan Quality Metric Component	Objective(s)	Result	Score	Max Score	Performance
[PTV_5040] V[50.4Gy] (%)	≥ 95 [> 90]	94.5513	29.33	30.00	97.8%
[PTV_5040] V[47.88Gy] (%)	≥ 100 [> 98]	99.9768	24.93	25.00	99.7%
[PTV_5040] V[52.92Gy] (%)	≤ 0 [< 10]	0.0232	9.98	10.00	99.8%
[PTV_5040] D[0.03cc] (Gy)	≤ 52 [< 55]	52.9810	6.73	10.00	67.3%
[PTV_5040] Conformation Number [47.88Gy]	≥ 1 [> 0.5]	0.5803	1.28	5.00	25.7%
[CTV] V[50.4Gy] (%)	≥ 99 [> 94]	99.3392	25.00	25.00	100.0%
[OPTIC CHIASM] D[0.03cc] (Gy)	≤ 35 [< 46]	41.3545	4.65	8.00	58.1%
[illegible]	≤ 1 [< 10]	4.1456	3.90	5.00	78.1%
[illegible]	≤ 27 [< 45]	16.0740	8.00	8.00	100.0%
[illegible]	≤ 1 [< 25]	5.7387	5.95	7.00	85.0%
[illegible]	≤ 1 [< 5]	3.9097	4.03	5.00	80.6%
[RT ORBIT] Mean dose (Gy)	≤ 0 [< 10]	2.5415	2.15	3.00	71.8%
[RT OPTIC NERVE] D[0.03cc] (Gy)	≤ 10 [< 20]	8.5762	3.00	3.00	100.0%
[RT LENS] D[0.03cc] (Gy)	≤ 0.5 [< 3]	2.0525	1.58	3.00	52.6%
[CORD] D[0.03cc] (Gy)	≤ 48	0.8092	0.00	0.00	100.0%
[BRAINSTEM] D[0.03cc] (Gy)	≤ 50.4	49.0009	0.00	0.00	100.0%
Global Max Location (ROI)	CTV [CTV; PTV_5040]	PTV_5040	2.00	3.00	66.7%
Total [17 Metrics]			**132.51**	**150.00**	**88.3%**

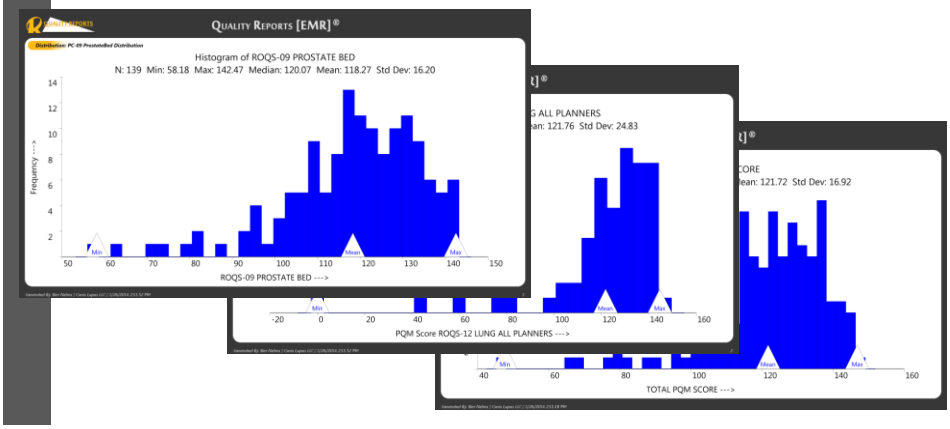


BUILDING A PERFORMANCE DISTRIBUTION



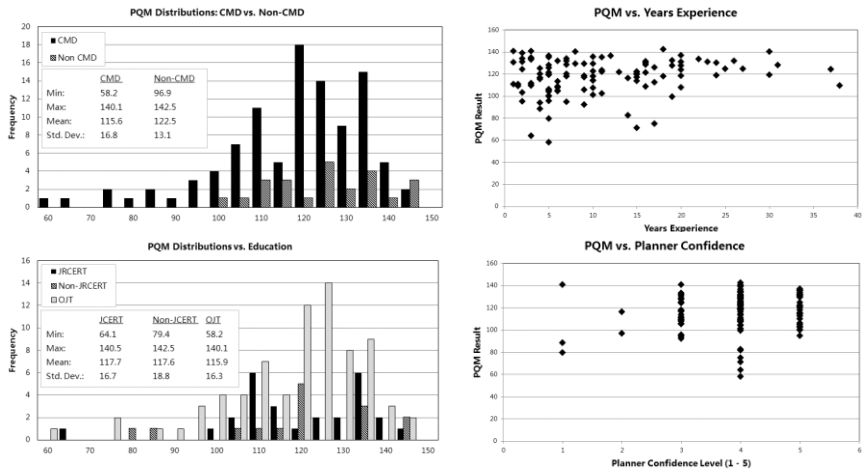
Some Plan Challenge Conclusions

- Despite controlled inputs (CT and structures) and well-defined objectives (Plan Quality Algorithm), ***THERE IS VERY HIGH VARIABILITY IN PLAN QUALITY***.



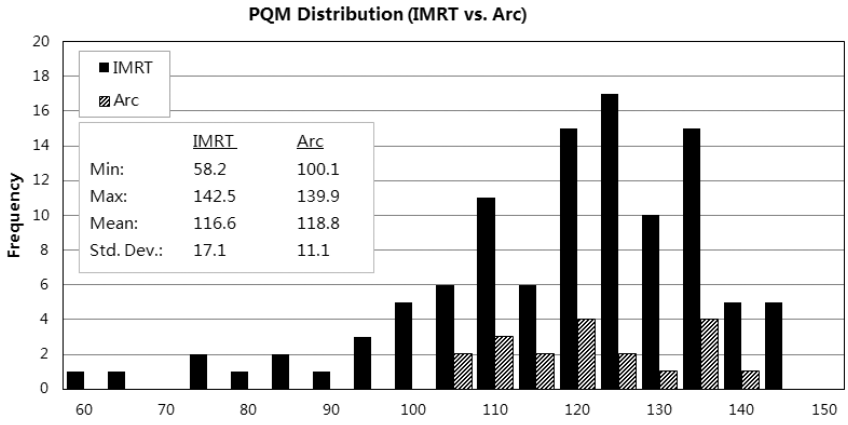
Some Plan Challenge Conclusions

- **NO CORRELATION WITH: CERTIFICATION, EDUCATION, EXPERIENCE, OR CONFIDENCE.**



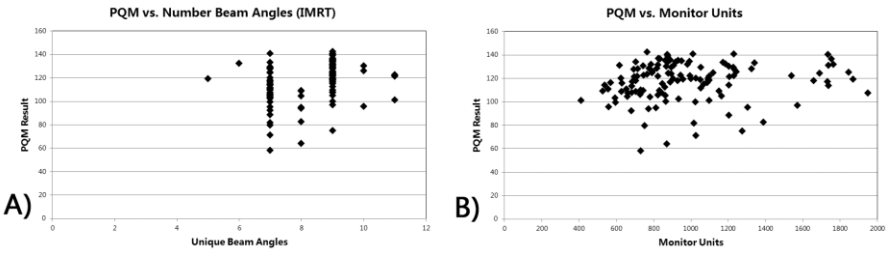
Some Plan Challenge Conclusions

- **VMAT is not better than IMRT (but it is less variable)**



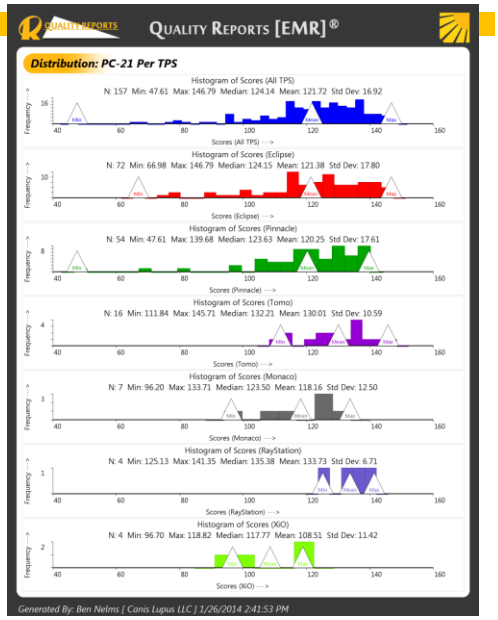
Some Plan Challenge Conclusions

- No correlation with plan complexity.
- **Plan Quality is an art, determined by skill level.**

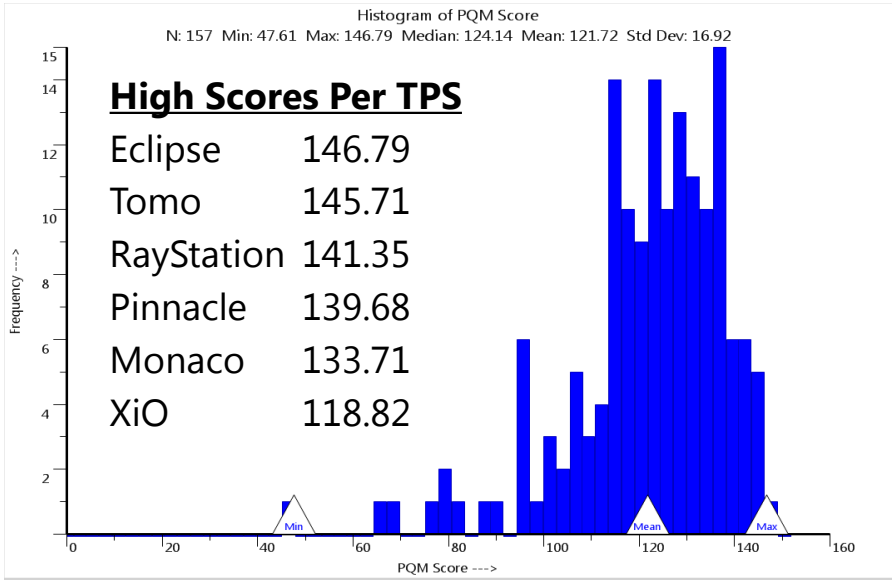


Some Plan Challenge Conclusions

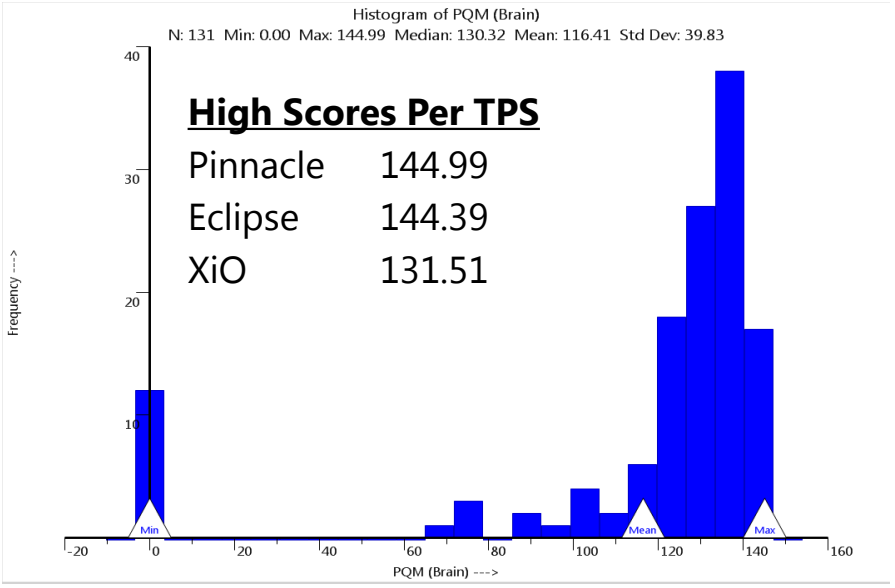
- **There is a dependence on TPS** (especially if considering the max potential of Plan Quality).



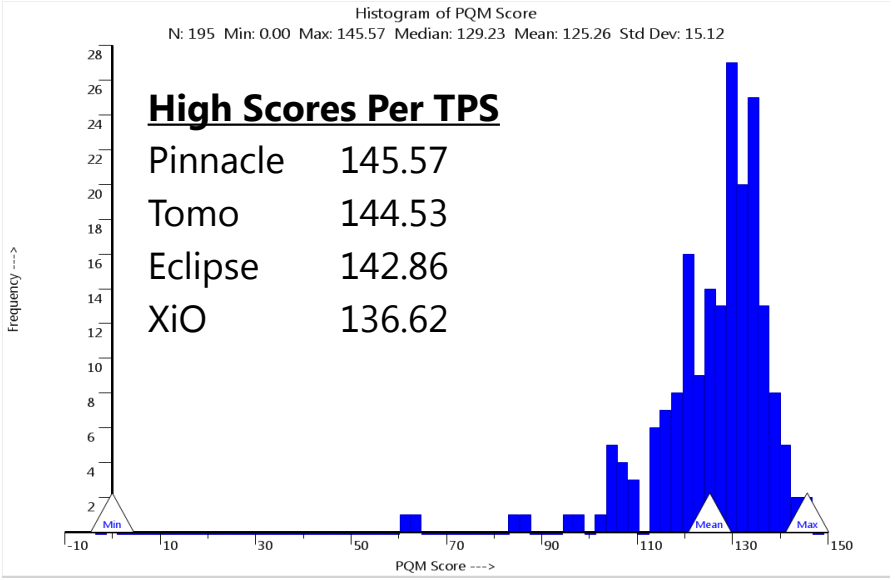
Plan Challenge [Anus]



PLAN CHALLENGE [BRAIN]



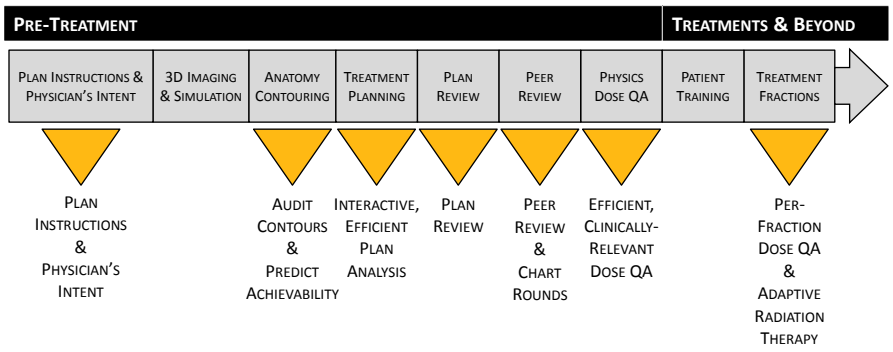
PLAN CHALLENGE [ABDOMEN]



PQM Applications

- **Clearly the Plan Quality Algorithm and the "PQM" results are a powerful way to define and measure Plan Quality.**
- **What are applications beyond general assessments of standard plans?**
  - Per-Patient Workflow
  - System Commissioning and Validation
  - Accreditation, Competency Testing, & Training

PQM Applications (Per Patient)



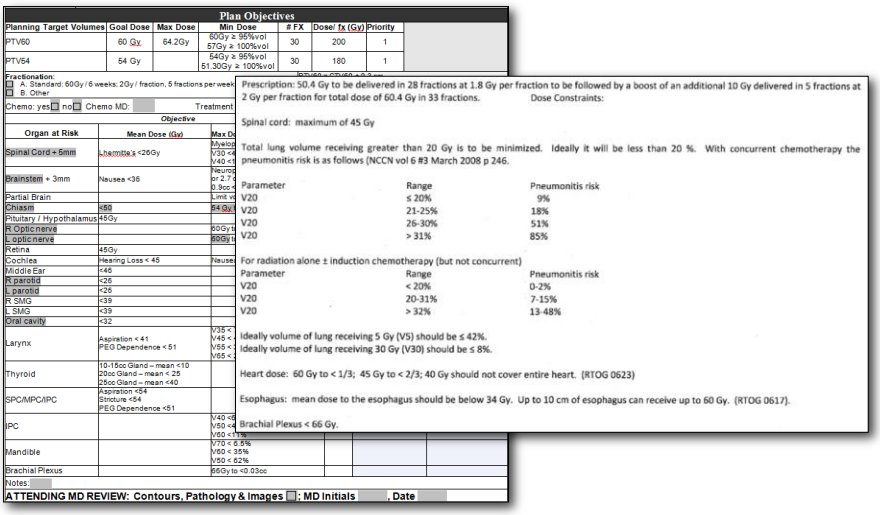
Plan Instructions & Physician's Intent

Plan Instructions & Physician's Intent	3D Imaging & Simulation	Anatomy Contouring	Treatment Planning	Plan Review	Peer Review	Physics Dose QA	Patient Training	Treatment Fractions

- **Required**
- **Must be traceable to what is eventually achieved**
- **Contains (but not limited to):**
  - Target prescription
  - Fractionation
  - OAR dose objectives
  - Required structures to be contoured
  - Physician's approval

Plan Instructions & Physician's Intent

- **Templates created per institution; highly variable**



Plan Objectives

Planning Target Volumes	Goal Dose	Max Dose	Min Dose	# FX	Dose/ fx (Gy)	Priority
PTV60	60 Gy	64.2Gy	60Gy ≥ 95%vol 57Gy ≥ 100%vol	30	200	1
PTV54	54 Gy		54Gy ≥ 95%vol 51.30Gy ≥ 100%vol	30	180	1

Fractionation:
☐ A. Standard: 60Gy / 6 weeks; 2Gy / fraction, 5 fractions per week
☐ B. Other

Chemo: yes☐ no☐ Chemo MD: Treatment

Objective

Organ at Risk	Mean Dose (Gy)	Max D
Spinal Cord + 5mm	Lhermitte's <26Gy	Myelop V30 <4 V40 <1
Brainstem + 3mm	Nausea <36	Neurop or 2.7 c 0.9cc >
Partial Brain		Limit vo
Chiasm	<50	54 Gy
Pituitary / Hypothalamus	45Gy	
R Optic nerve		60Gy t
L optic nerve		60Gy t
Retina	45Gy	
Cochlea	Hearing Loss < 45	Nause
Middle Ear	<45	
R parotid	<26	
L parotid	<26	
R SMG	<39	
L SMG	<39	
Oral cavity	<32	
Larynx	Aspiration < 41 PEG Dependence < 51	V35 < V45 < V55 < V65 <
Thyroid	10-15cc Gland – mean <10 20cc Gland – mean < 25 25cc Gland – mean <40	
SPC/MPC/IPC	Aspiration <54 Stricture <54 PEG Dependence <51	
IPC		V40 <6 V50 <4 V60 <1
Mandible		V70 < 6.5% V60 < 35% V50 < 62%
Brachial Plexus		66Gy to <0.03cc

Notes:

ATTENDING MD REVIEW: Contours, Pathology & Images ☐; MD Initials , Date

Prescription: 50.4 Gy to be delivered in 28 fractions at 1.8 Gy per fraction to be followed by a boost of an additional 10 Gy delivered in 5 fractions at 2 Gy per fraction for total dose of 60.4 Gy in 33 fractions.

Dose Constraints:

Spinal cord: maximum of 45 Gy

Total lung volume receiving greater than 20 Gy is to be minimized. Ideally it will be less than 20 %. With concurrent chemotherapy the pneumonitis risk is as follows (NCCN vol 6 #3 March 2008 p 246.

Parameter	Range	Pneumonitis risk
V20	≤ 20%	9%
V20	21-25%	18%
V20	26-30%	51%
V20	> 31%	85%

For radiation alone ± induction chemotherapy (but not concurrent)

Parameter	Range	Pneumonitis risk
V20	< 20%	0-2%
V20	20-31%	7-15%
V20	> 32%	13-48%

Ideally volume of lung receiving 5 Gy (V5) should be ≤ 42%.
Ideally volume of lung receiving 30 Gy (V30) should be ≤ 8%.

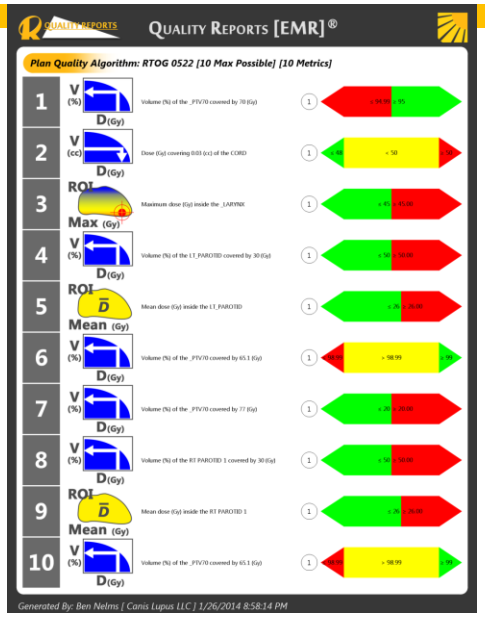
Heart dose: 60 Gy to < 1/3; 45 Gy to < 2/3; 40 Gy should not cover entire heart. (RTOG 0623)

Esophagus: mean dose to the esophagus should be below 34 Gy. Up to 10 cm of esophagus can receive up to 60 Gy. (RTOG 0617).

Brachial Plexus < 66 Gy.

Plan Instructions & Physician's Intent

- **Plan Quality Algorithm Report = Physician's Intent & Objectives**
  - Target and critical OAR goals are much more clearly documented in the Plan Quality Algorithm reports
  - Standardized
  - Accessible
  - Same data is used to generate the results ➔ guaranteed traceability



Audit Contours & Predict Achievability

Plan Instructions & Physician's Intent	3D Imaging & Simulation	Anatomy Contouring	Treatment Planning	Plan Review	Peer Review	Physics Dose QA	Patient Training	Treatment Fractions

- **Audit Contours**
  - All critical target volumes and OARs (including preferred naming conventions) are stored in the Plan Quality Algorithm.
  - You can run the algorithm post-contouring (and pre-planning or during planning) to ensure all required contours have been defined.
- **Pre-Plan Prediction of Achievability**
  - "Icarus" feature predicts achievability of dose objectives taking into account the unique patient anatomy
  - Allows for: 1) setting of realistic expectations and/or 2) adjustment of dose objectives that cannot be met
  - Discussed further in the "Research" section of this presentation

Interactive, Efficient Plan Review

Plan Instructions & Physician's Intent	3D Imaging & Simulation	Anatomy Contouring	Treatment Planning	Plan Review	Peer Review	Physics Dose QA	Patient Training	Treatment Fractions

- **Automated & Immediate Feedback of all vital objectives**
  - Target and OAR objectives assessed immediately and efficiently
  - Identifies failing metrics or areas for improvement
  - Feedback into optimization process
  - Guarantees that plan objectives are not a "moving target"; ideal and acceptable levels are clear
  - Mitigates risk of omission

Interactive, Efficient Plan Review

First Attempt

Plan Quality Metric Component	Objective(s)	Result	Score	Max Score	Performance
[PTV_68] V[68.0Gy] (%)	≥ 95 [> 92.99]	97.2777	30.00	30.00	100.0%
[PTV_68] D[0.03cc] (Gy)	≤ 71.4 [< 74.80]	75.0276	7.13	10.00	71.3%
[PTV_68] Volume of Regret [68.0Gy] (cc)	≤ 10 [< 40.00]	58.3935	0.00	10.00	0.0%
[PTV_68] Conformation Number [64.6Gy]	≥ 1 [> 0.599]	0.6266	2.08	5.00	41.5%
[PROSTATE_BED] V[68.0Gy] (%)	≥ 99 [> 96.99]	100.0000	10.00	10.00	100.0%
[PTV_56] V[56.0Gy] (%)	≥ 95 [> 92.99]	99.0450	30.00	30.00	100.0%
[PTV56 - PTV68] V[58.8Gy] (%)	≤ 5 [< 25.00]	65.2693	0.00	10.00	0.0%
Global Max Location (ROI)	PROSTATE_BED [PROSTATE_BED; PTV_68]	PROSTATE_BED	5.00	5.00	100.0%
[RECTUM] V[65.0Gy] (%)	≤ 5 [< 25.00]	11.3431	9.37	10.00	93.7%
[RECTUM] V[68.0Gy] (cc)	≤ 0 [< 5.001]	3.5390	6.48	10.00	64.8%
[RECTUM] V[40.0Gy] (%)	≤ 20 [< 45]	58.5779	0.00	10.00	0.0%
[RECTUM] Serial Slice Evaluation [34.0Gy]	PASS [PASS]	Fail	-10.00	0.00	0.0%
[POST_RECTUM] V[34.0Gy] (%)	≤ 30	15.0310	0.00	0.00	100.0%
[BLADDER] V[65.0Gy] (%)	≤ 15 [< 30.00]	18.4116	5.67	7.00	81.0%
[BLADDER] V[40.0Gy] (%)	≤ 40 [< 55.00]	57.1114	0.00	3.00	0.0%
Total [15 Metrics]			**95.73**	**150.00**	**63.8%**

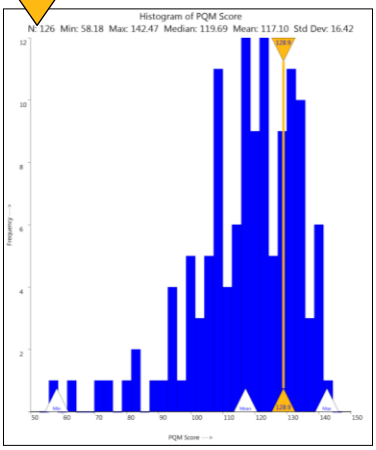
Final Plan

Score	Max Score	Performance
30.00	30.00	100.0%
9.50	10.00	95.0%
8.02	10.00	80.2%
4.01	5.00	80.1%
10.00	10.00	100.0%
30.00	30.00	100.0%
8.77	10.00	87.7%
3.00	5.00	60.0%
9.75	10.00	97.5%
8.44	10.00	84.4%
4.86	10.00	48.6%
0.00	0.00	100.0%
0.00	0.00	100.0%
5.14	7.00	73.4%
2.70	3.00	90.1%
134.18	**150.00**	**89.5%**

Peer Review & Chart Rounds

Plan Instructions & Physician's Intent	3D Imaging & Simulation	Anatomy Contouring	Treatment Planning	Plan Review	Peer Review	Physics Dose QA	Patient Training	Treatment Fractions

- **Peer Review & Chart Rounds**
  - Much more efficient because all the critical objectives and results are organized and scored
  - Clinical team is trained and vested in their Plan Quality Algorithms, so their peer reviews are standardized and effective
  - Compare each plan's performance vs. population of similar plans



Efficient & Relevant Per-Patient Dose QA

Plan Instructions & Physician's Intent	3D Imaging & Simulation	Anatomy Contouring	Treatment Planning	Plan Review	Peer Review	Physics Dose QA	Patient Training	Treatment Fractions

- **Per-Patient, Pre-Treatment Dose QA has evolved to use clinically-relevant metrics**
  - Plan Quality Algorithm can create PQM scoresheets that are much more efficient than per-metric analyses
  - Captures clinical impact as defined by the PQM metrics and priorities
  - **PQM_{Plan} ↔ PQM_{DoseQA}**:
    - Efficient
    - Comprehensive

Efficient & Relevant Per-Patient Dose QA

Results from Dose QA that estimates impact of TPS or delivery errors on patient dose

Plan Quality Metric Component	Objective(s)	Result	Score	Max Score	PLAN	DOSE QA
[PTV_5040] V[50.4Gy] (%)	≥ 95 [> 90]	95.1586	30.00	30.00	100.0%	
[PTV_5040] V[47.88Gy] (%)	≥ 100 [> 98]	99.8897	24.67	25.00	98.7%	
[PTV_5040] V[52.92Gy] (%)	≤ 0 [< 10]	0.0000	10.00	10.00	100.0%	
[PTV_5040] D[0.03cc] (Gy)	≤ 52 [< 55]	52.8440	7.19	10.00	71.9%	
[PTV_5040] Conformation Number [47.88Gy]	≥ 1 [> 0.5]	0.6924	3.08	5.00	61.6%	
[CTV] V[50.4Gy] (%)	≥ 99 [> 94]	98.5941	24.39	25.00	97.6%	
[OPTIC CHIASM] D[0.03cc] (Gy)	≤ 35 [< 46]	36.9924	7.20	8.00	90.0%	
[LT ORBIT] Mean dose (Gy)	≤ 1 [< 10]	1.7026	4.77	5.00	95.3%	
[LT OPTIC NERVE] D[0.03cc] (Gy)	≤ 27 [< 45]	21.5059	8.00	8.00	100.0%	
[LT COCHLEA] D[0.03cc] (Gy)	≤ 1 [< 25]	5.4529	6.01	7.00	85.9%	
[LT LENS] D[0.03cc] (Gy)	≤ 1 [< 5]	1.3503	4.88	5.00	97.7%	
[RT ORBIT] Mean dose (Gy)	≤ 0 [< 10]	0.9792	2.67	3.00	89.1%	
[RT OPTIC NERVE] D[0.03cc] (Gy)	≤ 10 [< 20]	13.9761	1.72	3.00	57.4%	
[RT LENS] D[0.03cc] (Gy)	≤ 0.5 [< 3]	0.8444	2.83	3.00	94.3%	
[CORD] D[0.03cc] (Gy)	≤ 48	5.6817	0.00	0.00	100.0%	
[BRAINSTEM] D[0.03cc] (Gy)	≤ 50.4	47.3644	0.00	0.00	100.0%	
Global Max Location (ROI)	CTV [CTV; PTV_5040]	PTV_5040	2.00	3.00	66.7%	
Total [17 Metrics]			**139.41**	**150.00**	**92.9%**	

Per-Fraction Dose QA & Adaptive RT

Plan Instructions & Physician's Intent	3D Imaging & Simulation	Anatomy Contouring	Treatment Planning	Plan Review	Peer Review	Physics Dose QA	Patient Training	Treatment Fractions

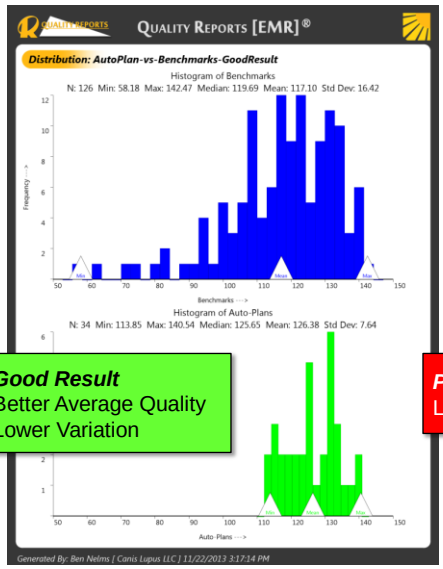
- **Per-Fraction**
  - Per-fraction needs to be as automated as possible to avoid being a high-inspection, resource drain
  - Run PQM results and set tolerances on score degradation levels, creating "red flag" events
- **Adaptive Radiation Therapy**
  - Cumulative dose accrued over all fractions analyzed with Plan Quality and $PQM_{Achieved}$ compared directly to $PQM_{Planned}$

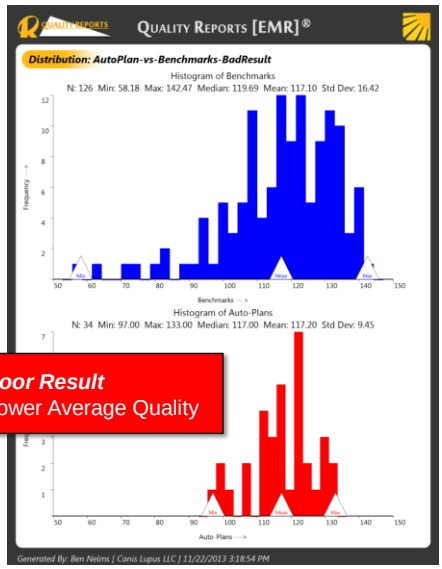
Other PQM Applications (Quality Management)

Commissioning, Validation			Accreditation, Competency Testing & Training		
TPS Commissioning	Machine Commissioning	TPS Upgrades & Version Testing	MOC Projects	Planner Assessment	Continuing Education
Validate Quality of TPS Optimization	Assess Deliverability as a Function of Plan Quality	Quantify Improvement (or lack thereof)	Assess Continual Improvement	Assess Planner Abilities	Identify Areas of Need, Per Planner

- **Evidence-Based / Comparative Effectiveness**
  - Sound strategy based on objective evidence
  - Statistical Process Control
  - Imperative in a "Pay-for-Performance" future

Comparative Effectiveness





Potential: Protocols & Clinical Trials

- **Easy audit of submitted plans (e.g. RTOG)**
  - Removes the high resource cost of generating metric results vs. goals
  - Removes the variability of methods
- **Retrospective analysis of different PQM algorithms vs. clinical outcomes**
  - Connect PQM Score with Outcomes.
  - Allows standardization of:
    - Plan Objectives
    - Plan Strategies
    - Plan Review Methods
    - Peer Review

Summary of PQM Applications