



# Outlining:effect on outcome and training

Sarah Gwynne  
Consultant Clinical Oncologist  
Singleton Hospital, Swansea  
NPL 13.12.13

# Outlining variation

- The weakest link in the RT process
- Inherently observer biased procedure
- First inter-observer variation study in RT reported by Kramer *et al* as early as 1977
- Typical reported magnitude of interclinician variation commonly exceeds that of geometric systematic error
- Big issue for clinical trials

Why does it matter?

# Why does it matter?

Because it can impact on outcome



Contents lists available at SciVerse ScienceDirect

## Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com



Systematic review

## QA makes a clinical trial stronger: Evidence-based medicine in radiation therapy

Damien C. Weber<sup>a,e,\*</sup>, Milan Tomsej<sup>b</sup>, Christos Melidis<sup>c</sup>, Coen W. Hurkmans<sup>d,e</sup>**Table 2**

Results of QART assessment with patient outcome in prospective clinical trials.

Study [ref]	Type of QA	Number of cases evaluated <i>n</i> (%)	Minor deviations <i>n</i> (%)	Major deviations <i>n</i> (%)	Technical issues with QA review <i>n</i> (%)	Impact on clinical outcome	<i>p</i> Value
HD 4 [5]	R	368 (98.0)	–	141 (37.5) <sup>*</sup>	8 (2.1)	7-year RFS with D: 72% vs. 7-year RFS with no D: 84%	0.004
EORTC 20884 [2]	R	135 (88.8)	–	63 (46.7)	46 (30.3)	5-year RFS with D: 90% vs. 5-year RFS without D: 84%	0.31
RTOG 0411 [4]	R	NS	–	13 (13.4)	NS	Grade GI ≥ 3 toxicity with D: 45% <sup>‡</sup> vs. Grade GI ≥ 3 toxicity without D: 18% <sup>‡</sup>	0.05
RTOG 9704 [1]	R	416 (92.2)	–	200 (48.0) <sup>**</sup>	14/35 (40.0) <sup>†</sup>	mOS with D: 1.46 yo vs. mOS without D: 1.74 yo	0.008
RTOG 0022 [8]	R	67 (97.0)	47 (89.0)	6 (11.0)	14/67 (21.0)	LRF with major D: 50% vs. LRF with no major D: 6%	0.04
TROG 0202 [15]	P & R <sup>††</sup>	687 (80.5) <sup>‡‡</sup>	–	97 (11.8)	33/820 (4.0)	OS with major D: 70% vs. OS without major D: 50%	<0.001 5

## Critical Impact of Radiotherapy Protocol Compliance and Quality in the Treatment of Advanced Head and Neck Cancer: Results From TROG 02.02

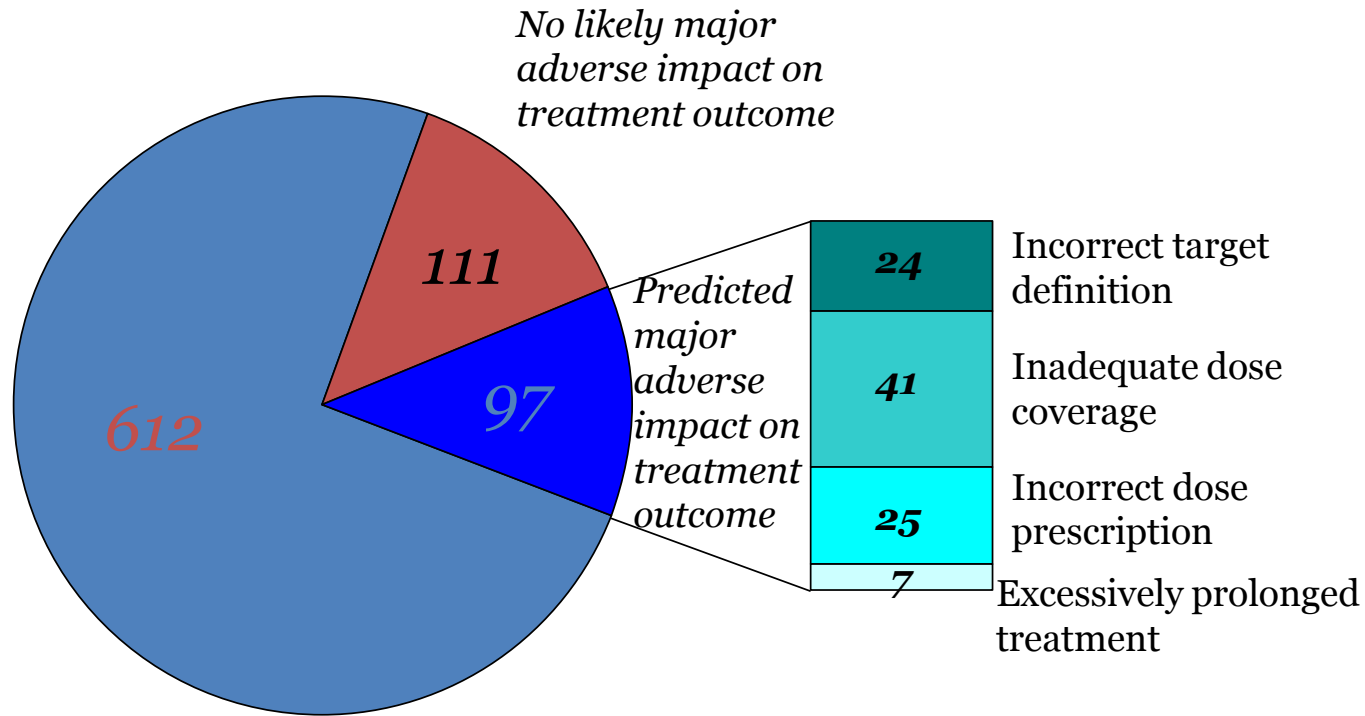
*Lester J. Peters, Brian O'Sullivan, Jordi Giralt, Thomas J. Fitzgerald, Andy Trotti, Jacques Bernier, Jean Bourhis, Kally Yuen, Richard Fisher, and Danny Rischin*

- **TROG 02.02 (Trans-Tasman Radiation Oncology Group) trial HeadSTART**
  - Tirapazamine, Cisplatin, and Radiation versus Cisplatin and Radiation for advanced SCC of the Head and Neck
  - RT data submitted by end of 1<sup>st</sup> week to Quality Assurance Review Centre (QARC)
  - After completion of treatment TMG review of all data (n=853)

# Critical Impact of Protocol Compliance

## Retrospective Review

***Non-Compliant (25.4%)***



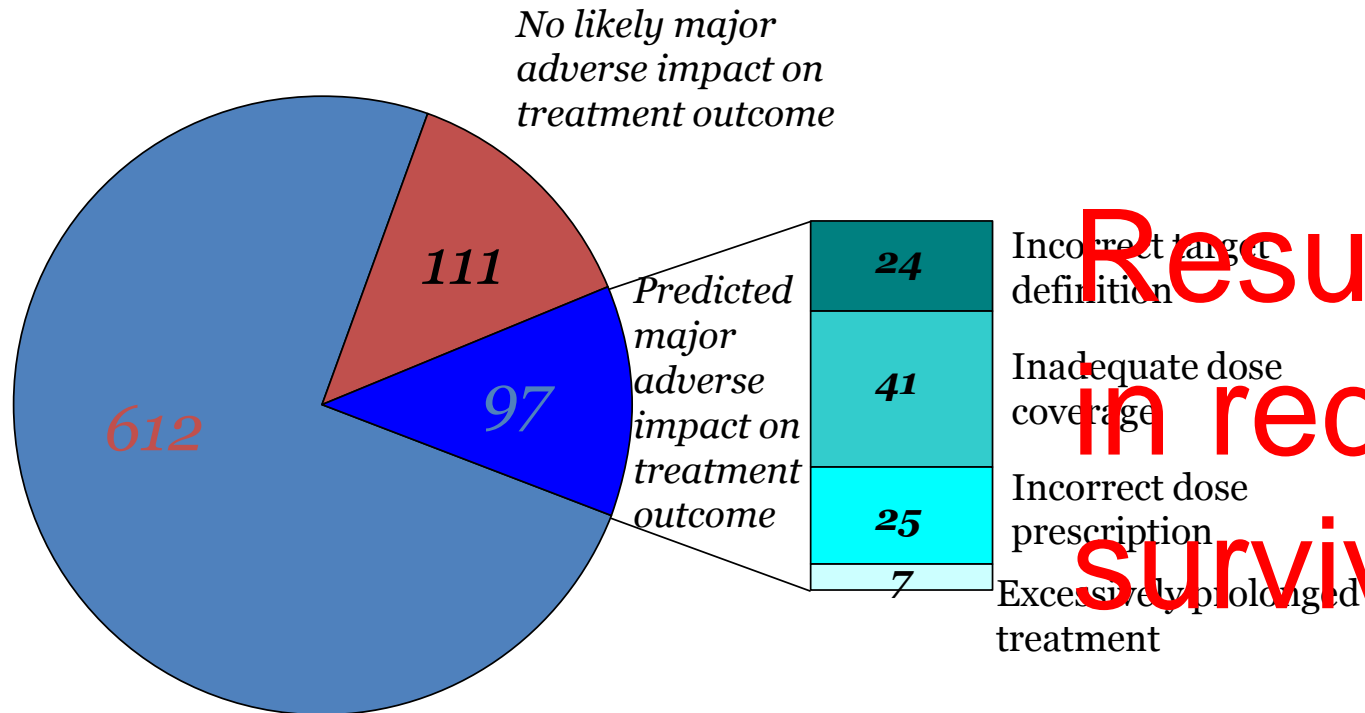
***Compliant (74.6%)***

***33 non-evaluable plans***

# Critical Impact of Protocol Compliance

## Retrospective Review

**Non-Compliant (25.4%)**



**Resulted in reduced survival**

**Compliant (74.6%)**

*33 non-evaluable plans*



RTOG guidelines

## Compliance with therapeutic guidelines in Radiation Therapy Oncology Group prospective gastrointestinal clinical trials <sup>☆</sup>

Christopher G. Willett <sup>a,\*</sup>, Jennifer Moughan <sup>b</sup>, Elizabeth O'Meara <sup>c</sup>, James M. Galvin <sup>c</sup>,  
 Christopher H. Crane <sup>d</sup>, Kathryn Winter <sup>b</sup>, Denise Manfredi <sup>c</sup>, Tyvin A. Rich <sup>e</sup>, Rachel Rabinovitch <sup>f</sup>,  
 Robert Lustig <sup>g</sup>, Mitchell Machtay <sup>h</sup>, Walter J. Curran <sup>i</sup>

**Table 2**

Multivariate analysis for overall survival: head of pancreas patients only ( $n = 359$ ) [31].

Adjustment variables	Comparison	Adjusted HR	$p$ -Value <sup>a</sup>
Treatment	Gemcitabine vs. 5-FU	0.79 (0.62–0.99)	0.043
Nodal involvement	No vs. Yes	1.47 (1.13–1.91)	0.0036
Tumor diameter	<3 vs. $\geq$ 3 cm	1.25 (0.98–1.59)	0.070
Surgical margin status	Negative	Ref. level	–
	Positive	1.07 (0.82–1.40)	0.64
	Unknown	0.94 (0.69–1.27)	0.68
RT QA score	<PP vs. PP	0.75 (0.60–0.95)	0.016

Abbreviations: 5-FU = fluorouracil; HR = hazard ratio; CI = confidence interval.

<sup>a</sup>  $p$  Value from chi-square test using the Cox proportional hazards model.

RTOG guidelines

## Compliance with therapeutic guidelines in Radiation Therapy Oncology Group prospective gastrointestinal clinical trials <sup>☆</sup>

Christopher G. Willett <sup>a,\*</sup>, Jennifer Moughan <sup>b</sup>, Elizabeth O'Meara <sup>c</sup>, James M. Galvin <sup>c</sup>,  
 Christopher H. Crane <sup>d</sup>, Kathryn Winter <sup>b</sup>, Denise Manfredi <sup>c</sup>, Tyvin A. Rich <sup>e</sup>, Rachel Rabinovitch <sup>f</sup>,  
 Robert Lustig <sup>g</sup>, Mitchell Machtay <sup>h</sup>, Walter J. Curran <sup>i</sup>

**Table 2**

Multivariate analysis for overall survival: head of pancreas patients only ( $n = 359$ ) [31].

Adjustment variables	Comparison	Adjusted HR	$p$ -Value <sup>a</sup>
Treatment	Gemcitabine vs. 5-FU	0.79 (0.62–0.99)	0.043
Nodal involvement	No vs. Yes	1.47 (1.13–1.91)	0.0036
Tumor diameter	<3 vs. $\geq$ 3 cm	1.25 (0.98–1.59)	0.070
Surgical margin status	Negative	Ref. level	–
	Positive	1.07 (0.82–1.40)	0.64
	Unknown	0.94 (0.69–1.27)	0.68
RT QA score	<PP vs. PP	0.75 (0.60–0.95)	0.016

Abbreviations: 5-FU = fluorouracil; HR = hazard ratio; CI = confidence interval.

<sup>a</sup>  $p$  Value from chi-square test using the Cox proportional hazards model.

Phase II Study of Bevacizumab With Concurrent  
Capecitabine and Radiation Followed by Maintenance  
Gemcitabine and Bevacizumab for Locally Advanced  
Pancreatic Cancer: Radiation Therapy Oncology Group  
RTOG 0411

*Christopher H. Crane, Kathryn Winter, William F. Regine, Howard Safran, Tyvin A. Rich, Walter Curran,  
Robert A. Wolff, and Christopher G. Willett*

from the Department of Radiation  
Oncology and Gastrointestinal Medical

Embedded QA programme with retrospective review of GTV outlining

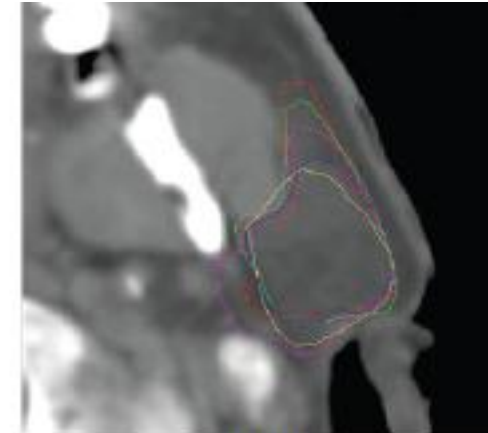
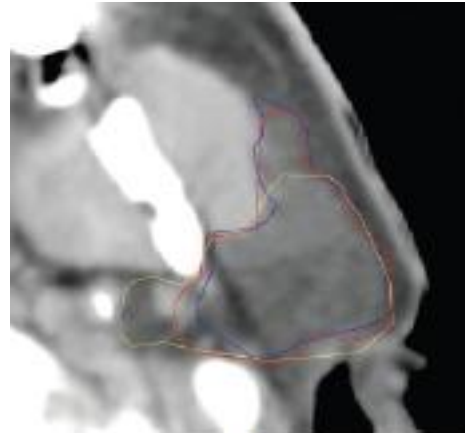
11 (13.4%) unacceptable deviations in outlining (defined by GTV >5cm than  
size of tumour, no GTV delineated)

Significant correlation between major deviation and incidence of  $\geq$ grade 3 GI  
toxicity in both the chemoradiotherapy (45 vs 18%,  $p=0.05$ ) and maintenance  
chemotherapy (45% vs 13%  $p=0.01$ ) components of treatment

## **Interobserver variation in parotid gland delineation: a study of its impact on intensity-modulated radiotherapy solutions with a systematic review of the literature**

S W LOO, FRCR, W M C MARTIN, FRCR, P SMITH, BSc, S CHERIAN, FRCR and T W ROQUES, FRCR

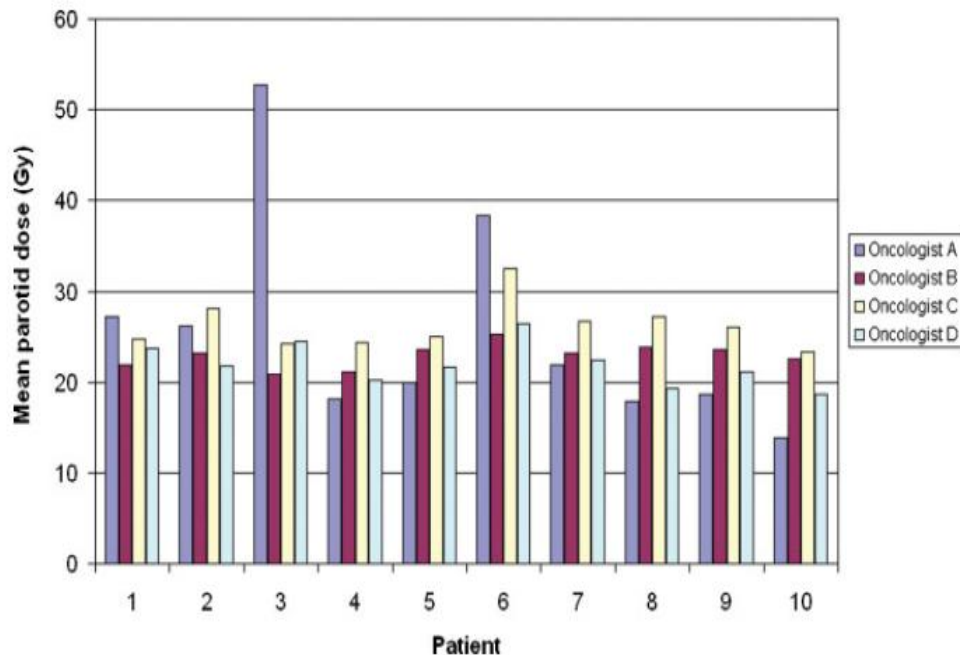
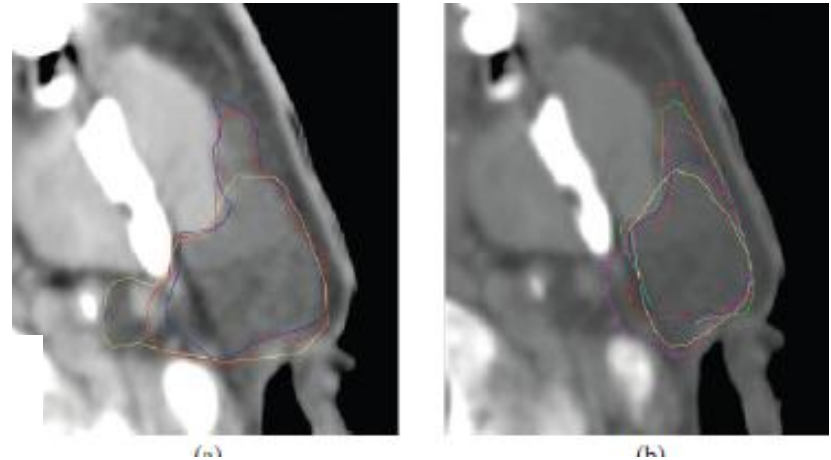
10 cases of stage 4 oropharynx  
Parotids outlined by 4 rad oncs and 3 radiologists  
Target to C/L parotid was 10% of 24Gy



## Interobserver variation in parotid gland delineation: a study of its impact on intensity-modulated radiotherapy solutions with a systematic review of the literature

S W LOO, FRCR, W M C MARTIN, FRCR, P SMITH, BSc, S CHERIAN, FRCR and T W ROQUES, FRCR

10 cases of stage 4 oropharynx  
 Parotids outlined by 4 rad oncs and 3 radiologists  
 Target to C/L parotid was 10% of 24Gy



During actual treatment dose delivered to parotid was within 10% of 24Gy in all cases

In this study this target achieved by 53% of volumes delineated by radiation oncologists

Parotid DVHs of 46% of the study contours would have resulted in a different IMRT plan

What to do about it

# Causes of variation

- Several identified
- Two main categories
  - Imaging
  - Observers

# Reducing interobserver variation

- Variety of interventions (trial and non-trial)
- Imaging
  - Better imaging
- Observers
  - Training
  - Clear protocol and access to an outlining atlas
  - Outlining of pre-trial benchmark cases/review of clinical cases from each centre (“dummy runs”)
  - Workshop attendance
  - Real time review of outlining



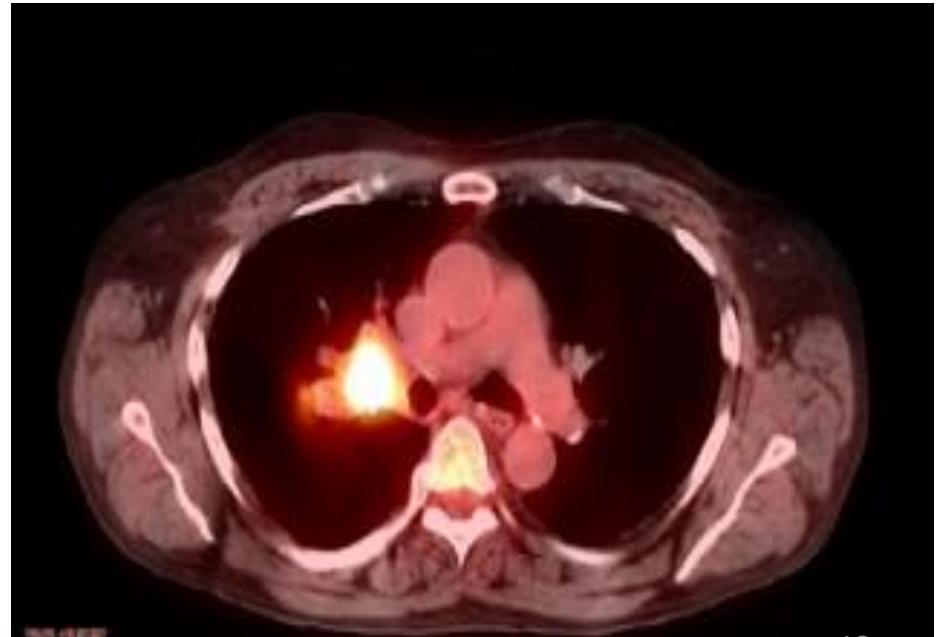
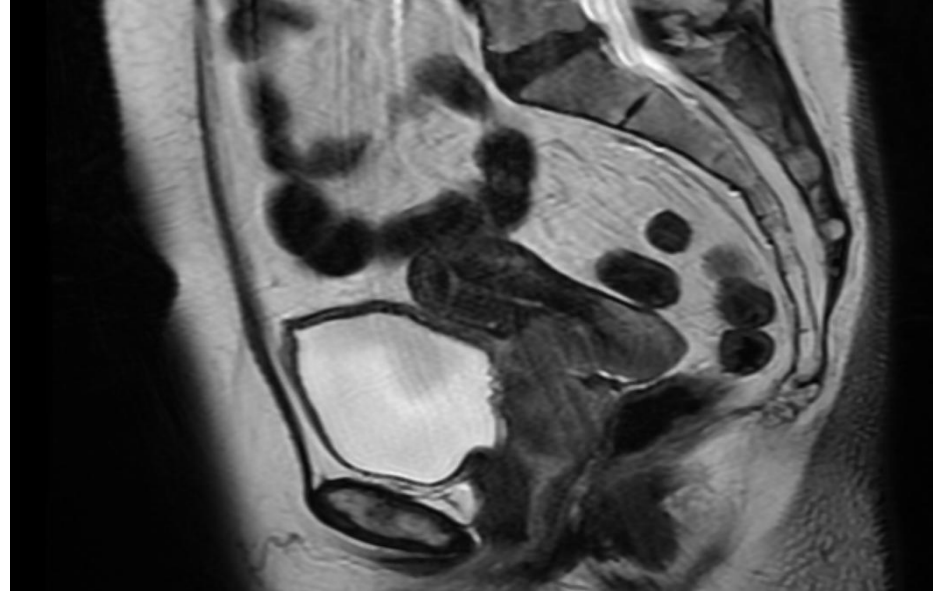
# What works?

(v little randomised evidence to confirm)

- Access to a protocol and an outlining atlas improves consistency in outlining in prostate and rectal cancers, respectively
- Educational sessions or workshops reduce TVD variation in a range of settings such as prostate and lung
- In RTOG GI trials pre-treatment review has increased compliance

# Begin at the beginning..

- Use best imaging for TVD
- Contrast where appropriate
- Co-register modalities where appropriate
- Correct windowing
- Use multiple orthogonal views for TVD
- Involve radiology



RESEARCH

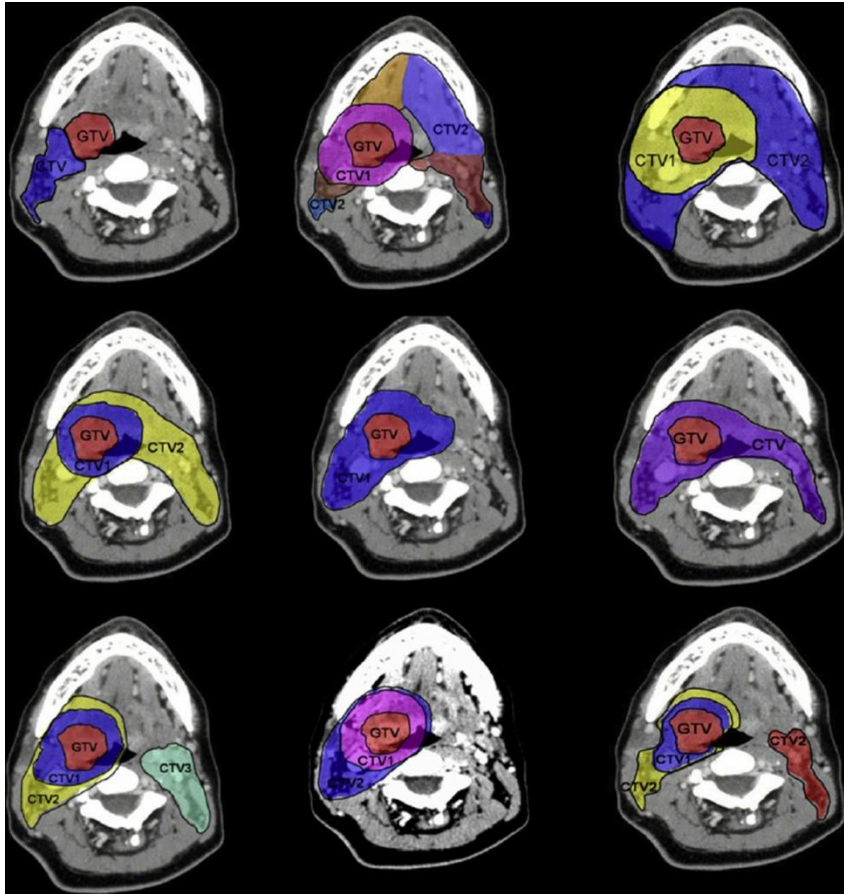
Open Access

# Decreased 3D observer variation with matched CT-MRI, for target delineation in Nasopharynx cancer

Coen RN Rasch<sup>1\*</sup>, Roel JHM Steenbakkers<sup>2</sup>, Isabelle Fitton<sup>3</sup>, Joop C Duppen<sup>1</sup>, Peter JCM Nowak<sup>4</sup>, Frank A Pameijer<sup>5</sup>, Avraham Eisbruch<sup>6</sup>, Johannes HAM Kaanders<sup>7</sup>, Frank Paulsen<sup>8</sup>, Marcel van Herk<sup>1</sup>

- Biggest differences seen at the top and bottom
- Using additional orthogonal views gives better concordance
- Better imaging improves consistency

# Clear protocol for outlining



- Heterogeneity in H&N target delineation. Nine distinct CTV designs which illustrate broad practitioner-dependent variation in target delineation strategies for the identical tonsil cancer case.

Hong *et al* Radioth Onc 2012

# Atlases

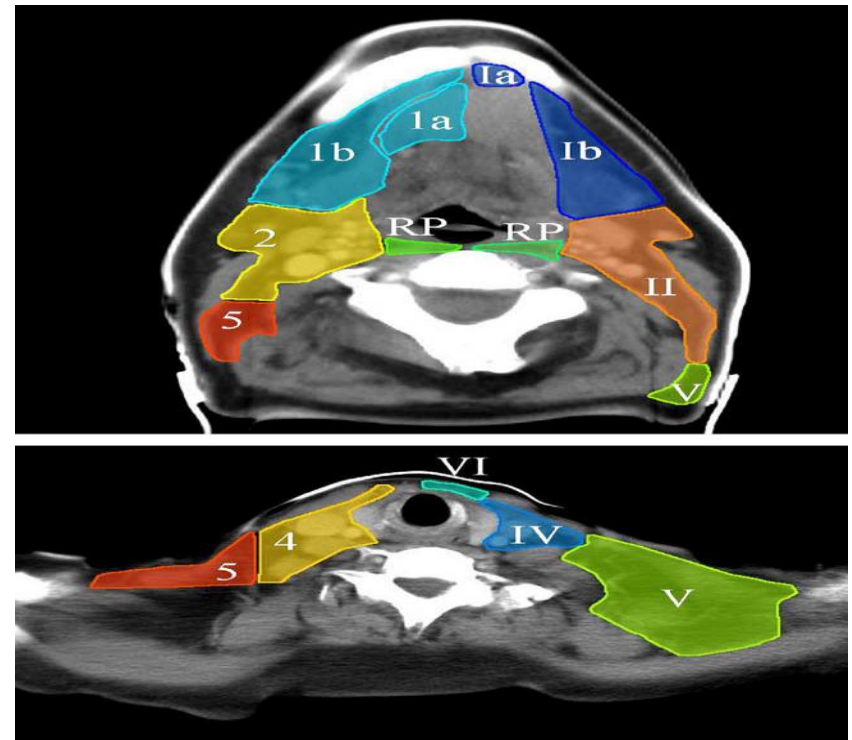
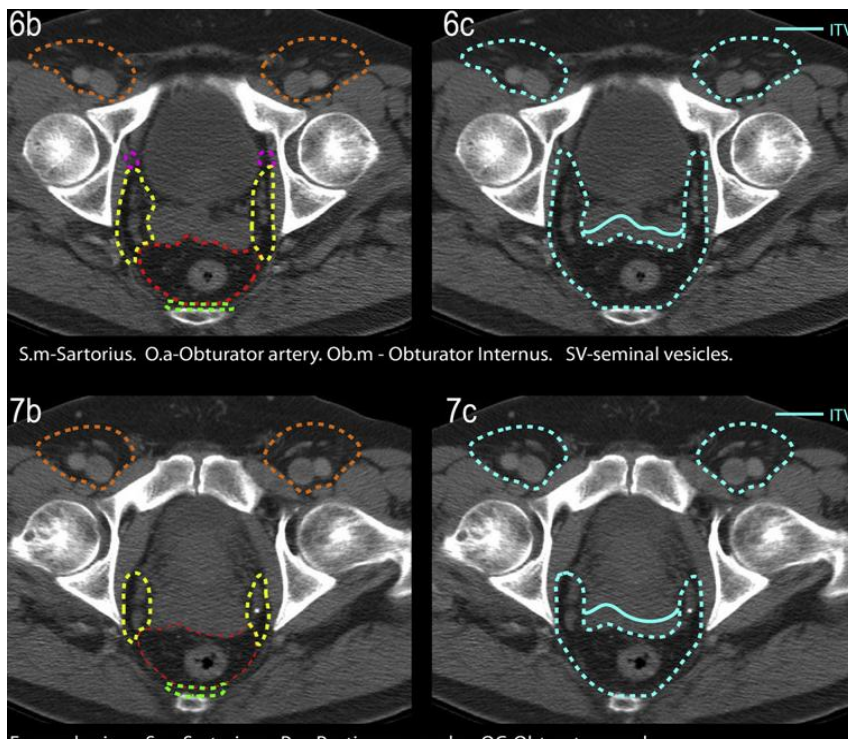
Evidence based approach to improving concordance (along with training sessions)

**Anal IMRT contouring atlas,  
AGITG**

**IJROBP 2012**

**Neck lymph node CTVs,  
Gregoire**

**Rad Onc 2003**



# NCRI RTQA OISG guidelines (based on RTQA experience in GI trials)

Cite this article as:

Gwynne S, Spezi E, Sebag-Montefiore D, Mukherjee S, Miles E, Conibear J, et al. Improving radiotherapy quality assurance in clinical trials: assessment of target volume delineation of the pre-accrual benchmark case. *Br J Radiol* 2013;86:20120398.

## REVIEW ARTICLE

# Improving radiotherapy quality assurance in clinical trials: assessment of target volume delineation of the pre-accrual benchmark case

<sup>1,2</sup>S GWYNNE, MD, FRCR, <sup>3</sup>E SPEZI, BSc, PhD, <sup>4</sup>D SEBAG-MONTEFIORE, FRCR, FRCR,  
<sup>5</sup>S MUKHERJEE, MD, FRCR, <sup>6</sup>E MILES, DCR(T), MPhil, <sup>6</sup>J CONIBEAR, MSc, FRCR, and  
<sup>7</sup>J STAFFURTH, MD, FRCR, on behalf of the NCRI RTQA OUTLINING AND IMAGING SUBGROUP

# NeosCOPE

## example of trial outlining QA

- Clear protocol (tested and revised after feedback from users)
- Atlas developed to help with delineation
- Workshop to discuss protocol and cases
- Pre-trial outlining assessment
- On-trial 'real time' review of 1<sup>st</sup> case from each centre



**Clinical  
Oncology**

**The Royal College of Radiologists**

## **Radiotherapy Planning – New Challenges for Clinical Oncologists**

*Raising standards in radiotherapy – keeping pace with new techniques*

More time to outline  
More training  
More quality assurance

Royal College of Radiologist Position Paper 2012  
<https://www.rcr.ac.uk/content.aspx?PageID=2069>



# Training



[Home](#) / [Clinical Oncology](#) / [Training and qualifications](#) / [Specialty training](#) / [COAST](#) / [Webcast 1 : Introduction to COAST, loading a Datapack and reviewing images](#)

[Webcast 1 : Introduction to COAST, loading a Datapack and reviewing images](#)

Please note that compatibility mode needs to be enabled to view this video in IE8 or IE9.



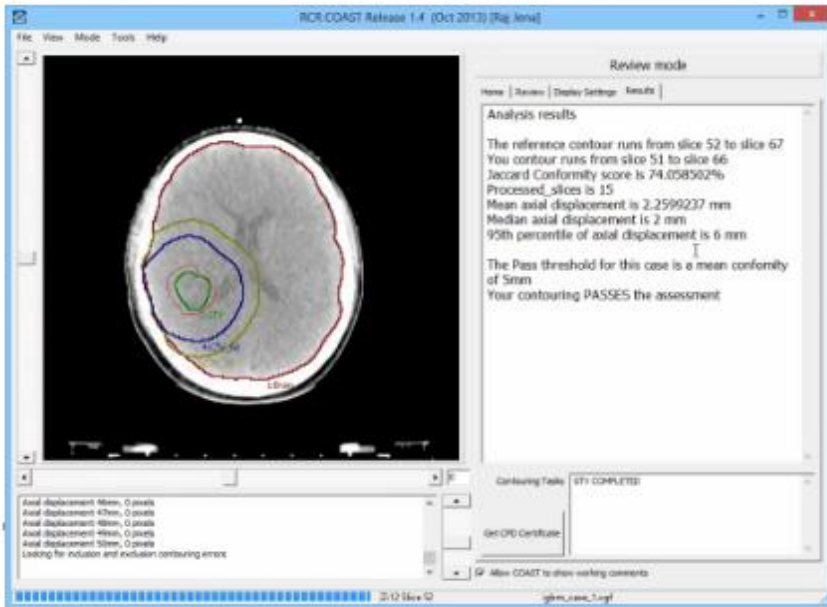
(Fellowship in Anatomic delineation and CONtouring)



# COAST

## Webcast 4 : How to score your work, and review contouring

Please note that compatibility mode needs to be enabled to view this video in IE8 or IE9.

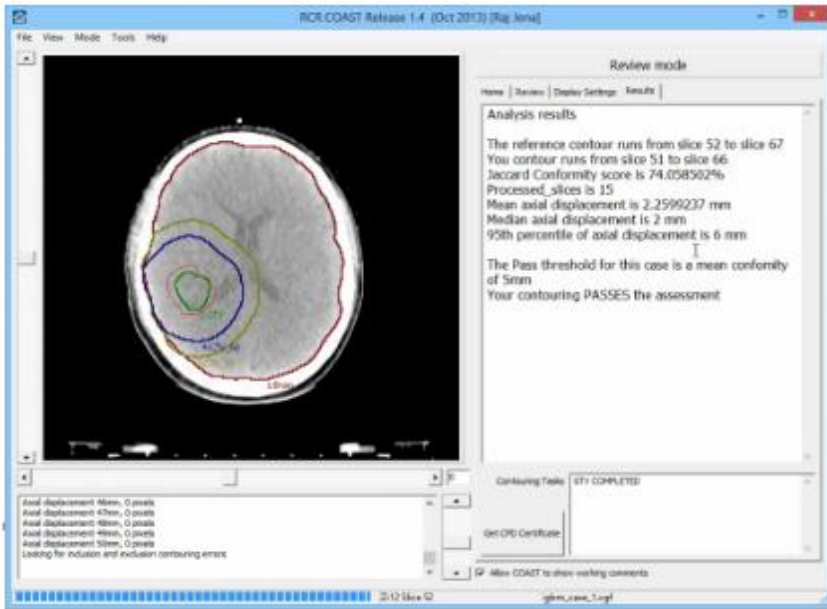


- Parotid (ART DECO)
- Glioma with MRI
- Mid 1/3 oesophagus (SCOPE 1)
- Lower 1/3 oesophagus (NeoSCOPE)
- Non small cell lung cancer
- Prostate and seminal vesicles

# COAST

## Webcast 4 : How to score your work, and review contouring

Please note that compatibility mode needs to be enabled to view this video in IE8 or IE9.



- Parotid (ART DECO)
- Glioma with MRI
- Mid 1/3 oesophagus (SCOPE 1)
- Lower 1/3 oesophagus (NeoSCOPE)
- Non small cell lung cancer
- Prostate and seminal vesicles

**Future role in revalidation of clinical oncologists - TBC**

# Conclusion

- Outlining variation exists and it matters
- GTV and OARs both important in 3D era
- Need to reduce variation as much as possible
- Training for all situations, additional trial specific training
- Clear protocols and atlases for trials
- Need for good quality QA – retrospective and increasing need for prospective
- Planning also critical!