[Article Full Title]
In Regard to Shortall et al

[Short Running Title]
In Regard to Shortall et al

[Author Names]
Martin A Ebert[1,2,3], Marco Marcello[1], Angel Kennedy[1], Annette Haworth[4], Lois C Holloway[4,5,6,7], Peter Greer[8,9], Jason A Dowling[4,6,7,8,10], Michael G Jameson[11,12], Dale Roach[6], David J Joseph[3,13,14], Sarah L Gulliford[15,16], Matthew R Sydes[17], Emma Hall[18], David P Dearnaley[19]

[Author Institutions]
1. Department of Radiation Oncology, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia
2. Department of Physics, University of Western Australia, Crawley, Western Australia, Australia
3. 5D Clinics, Claremont, Western Australia, Australia
4. School of Physics, University of Sydney, Camperdown, New South Wales, Australia
5. Department of Medical Physics, Liverpool Cancer Centre, Liverpool, New South Wales, Australia
6. South Western Sydney Clinical School, University of New South Wales, Liverpool, New South Wales, Australia
7. Centre for Medical Radiation Physics, University of Wollongong, Wollongong, New South Wales, Australia
8. School of Mathematical and Physical Sciences, University of Newcastle, Callaghan, New South Wales, Australia
9. Department of Radiation Oncology, Calvary Mater Newcastle, Waratah, New South Wales, Australia
10. CSIRO, Herston, Queensland, Australia
11. GenesisCare, Alexandria, New South Wales, Australia
12. St Vincent's Clinical School, University of New South Wales, Sydney, Australia
13. GenesisCare WA, Wembley, Western Australia, Australia
14. School of Surgery, University of Western Australia, Crawley, Western Australia, Australia
15. Radiotherapy Department, University College London Hospitals NHS Foundation Trust, London, United Kingdom
16. Department of Medical Physics and Biomedical Engineering, University College London, London, United Kingdom
17. MRC Clinical Trials Unit, Institute of Clinical Trials and Methodology, University College, London, London, United Kingdom
18. Clinical Trials and Statistics Unit, The Institute of Cancer Research, Sutton, United Kingdom
19. Academic UroOncology Unit, The Institute of Cancer Research and the Royal Marsden NHS Trust, London, United Kingdom

[Corresponding Author Name & Email Address]

Martin A Ebert
Martin.Ebert@health.wa.gov.au

[Author Responsible for Statistical Analysis Name & Email Address]
N/A

[Conflict of Interest Statement for All Authors]
Conflict of Interest:
Martin A Ebert: None
Marco Marcello: None
Angel Kennedy: None
Annette Haworth: None
Lois C Holloway: None
Peter Greer: None
Jason A Dowling: None
Michael G Jameson: None
Dale Roach: None
David J Joseph: None
Sarah L Gulliford: None
Matthew R Sydes:

- *Astellas* — *Research grants and drug costs or biomarker testing costs, all to institution and all active in past 36 months but on research outside of this research*
- *Clovis Oncology* — *Research grants and biomarker testing costs, all to institution and all active in past 36 months but on research outside of this research*
- *Janssen* — *Research grants and drug costs or biomarker testing costs, all to institution and all active in past 36 months but on research outside of this research*
- *Pfizer* — *Research grants and biomarker testing costs, all to institution and all active in past 36 months but on research outside of this research*
- *Novartis* — *Research grants and biomarker testing costs, all to institution and all active in past 36 months but on research outside of this research*
- *Sanofi-Aventis* — *Research grants and biomarker testing costs, all to institution and all active in past 36 months but on research outside of this research*
- *Lilly Oncology* — *Speaker fees at clinical trial statistics training meeting for clinicians (no discussion of particular drugs)*
- *Janssen* — *Speaker fees at clinical trial statistics training meeting for clinicians (no discussion of particular drugs)*
- *Independent member of many Independent Data Monitoring Committees but all for academic sponsors and none paid*

Emma Hall:

- *Cancer Research UK*	- *E Hall grantholder for Clinical Trials Project Grant to host institution (ICR)to support central trial running costs at ICR-CTSU.* - *Clinical Trials Unit programmatic grant funding to host institution to provide core support to ICR-CTSU.*
- *Accuray Inc.*	- *Grant received by Institution as contribution to the central trial management of a radiotherapy trial in prostate cancer.*
- *Varian Medical Systems Inc.*	- *Grant received by Institution as contribution to the central trial management of a radiotherapy trial in prostate cancer.*
- *Astra Zeneca*	- *Grant received by Institution as contribution to support central trial costs at ICR-CTSU for academic trials in prostate cancer.* - *Provision of study drug & placebo*
- *Janssen-Cilag*	- *Grant received by Institution as contributionto support trial health economic evaluation in an academic trial in prostate cancer.*
- *Bayer*	- *Grant received by Institution as contribution to support central trial costs at ICR-CTSU for an academic trial in prostate cancer.* - *Provision of study drug*
- *Merck Sharp & Dohm*	- *Grant received by Institution as contribution to support central trial costs at ICR-CTSU for an academic trial in prostate cancer*

David P Dearnaley:

- *Cancer Research UK*	*Programme Grants C46/A2131 (1999-2004), C46/A3976 (2004-2009), C46/A10588 (2009-14), C33589/A19727 (2015-19)*
- *Institute of Cancer Research*	*My previous employer, The Institute of Cancer Research receives loyalty income from abiraterone. I receive a share of this income through the ICR's Rewards to Discoverer's Scheme*
- *Location and stabilisation device*	*European Patent issued EP1933709B1. Pending in Canada and India.*
- *Jansen Advisory Boards*	*2015-2021. Consultancy fees*

[Funding Statement]
Funding: None.

[Data Availability Statement for this Work]
Not applicable

[Acknowledgements]
Not applicable

Shortall et al [1] have emphasised the need for statistical rigour in the development of voxel-level outcomes analyses and implied that three relevant published studies from Marcello et al [2-4] have not appropriately managed the multiple comparisons problem. “Statistical rigour” should be relative to an investigation’s objective. The exploratory studies of Marcello et al were undertaken with an awareness of the issues associated with the development of models spanning ~10^6 voxel-level features over cohorts with narrow variations in irradiation technique. An inflation in Type-I error had to be balanced against the chance of ignoring a potentially hypothesis-generating discovery. To achieve this balance, Marcello et al utilised several complementary methods:

1. Examining the distribution of the maximum test statistic (normalized dose difference) across the full distribution over permutations of the considered event labels, applying the methodology as reported in Chen et al [5] to generate adjusted p-values. Unfortunately, Shortall et al have incorrectly reported that Marcello et al used per-voxel permutation testing based explicitly on mean dose difference.
2. Independent voxel-level Cox proportional hazards modelling presented as hazard ratios with unadjusted p-values. Given the ambiguity regarding the number and variety of such exploratory analyses and difficulty in estimating the familywise error rate [6], the presentation of p-values at all is questionable. Marcello et al preferred to present hazard ratios and unadjusted p-values, with full associated disclosure. We draw attention to related studies that have singled out individual voxels as model variables with a priori interest, such as that by Witte et al [7]. Marcello have presented models for all such voxels and have highlighted the inevitability of false-positive results.
3. Assessment with multiple independent measurements. A much less elaborate contrivance can demonstrate the impact of multiple comparisons than the study of Bennet et al [8] – one needs to just randomly allocate event labels in their own dataset, or to just apply statistical reasoning. By assessing three cohorts, representing diversity in technique, populations and investigator teams, Marcello et al have assessed the preservation of associations across independent measurements.

The least absolute shrinkage and selection operator (LASSO) method was applied to reduce the number of features in multivariate models. We agree with Shortall et al that implying this as a method to account for multiple comparisons may have been misleading. We also agree that the resulting isolation of scattered voxels was not particularly helpful in this analysis, and highlight that alternative methods that accommodate spatial correlation are available, as recently summarised in Ebert et al [9].

Shortall et al also suggest that the structure-based Dice index could have been used to assess regional spatial mismatch in Marcello et al’s method. However, to avoid observer bias and pathological assumptions, Marcello et al deliberately took an anatomy/structure-agnostic approach. Marcello et al chose to assess robustness by comparing results in a template geometry that showed the largest difference relative to the selected common reference, with both identified via similarity-clustering [10]. Without objective ground-truth information and potentially systematic uncertainty introduced via a single template, we do not see justification for simply “blurring the dose distribution” (Shortall et al).

Shortall et al identify that significant model variables aside from dose had not been stated: these were included in the appendices to the three publications of Marcello et al., and included in models as described in the papers.

We are very enthusiastic to see this field progress and the diligence of Shortall et al is welcome. We hope though that their ambitions to be rigid and prescriptive will not stifle exploration in this area, and not impede fortuitous discovery at a time when the field is still developing.

1. Shortall, J., et al., *Flogging a Dead Salmon? Reduced Dose Posterior to Prostate Correlates With Increased PSA Progression in Voxel-Based Analysis of 3 Randomized Phase 3 Trials.* Int J Radiat Oncol Biol Phys, 2021. **110**: p. 696-699.
2. Marcello, M., et al., *Reduced Dose Posterior to Prostate Correlates With Increased PSA Progression in Voxel-Based Analysis of 3 Randomized Phase 3 Trials.* International Journal of Radiation Oncology, Biology, Physics, 2020. **108**: p. 1304-1318.
3. Marcello, M., et al., *Relationships between rectal and perirectal doses and rectal bleeding or tenesmus in pooled voxel-based analysis of 3 randomised phase III trials.* Radiotherapy and Oncology, 2020. **150**: p. 281-292.
4. Marcello, M., et al., *Increased Dose to Organs in Urinary Tract Associates With Measures of Genitourinary Toxicity in Pooled Voxel-Based Analysis of 3 Randomized Phase III Trials.* Frontiers in Oncology, 2020. **10**
5. Chen, C., M. Witte, W. Heemsbergen, and M. van Herk, *Multiple comparisons permutation test for image based data mining in radiotherapy.* Radiat Oncol, 2013. **8**: p. 293.
6. Rubin, M., *Do p values lose their meaning in exploratory analyses? It depends how you define the familywise error rate.* Review of General Psychology, 2017. **21**: p. 269-275.
7. Witte, M.G., et al., *Relating dose outside the prostate with freedom from failure in the Dutch trial 68 Gy vs. 78 Gy.* Int J Radiat Oncol Biol Phys, 2010. **77**: p. 131-8.
8. Bennett, C., M. Miller, and G. Wolford, *Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction.* Neuroimage, 2009. **47**: p. S125.
9. Ebert, M.A., et al., *Spatial descriptions of radiotherapy dose: normal tissue complication models and statistical associations.* Physics in Medicine & Biology, 2021. **66**: p. 12TR01.
10. Kennedy, A., et al., *Similarity clustering-based atlas selection for pelvic CT image segmentation.* Med Phys, 2019. **46**: p. 2243-2250.