SUPPLEMENTARY MATERIAL

Co-occurrence of symptoms after radiochemotherapy in locally advanced cervix cancer patients: a cluster analysis.

Marta Pelizzola[a*], Kari Tanderup[a,b], Supriya Chopra[c], Ina M. Jürgenliemk-Schulz[d], Remi Nout[e], Kathrin Kirchheiner[f] and Sofia Spampinato[a]

[a] Danish Center for Particle Therapy, Aarhus University Hospital, 8200 Aarhus, Denmark
[b] Department of Clinical Medicine, Aarhus University, 8200 Aarhus, Denmark
[c] Department of Radiation Oncology, Advanced Centre for Treatment Research and Education in Cancer, Tata Memorial Centre, Homi Bhabha National Institute, Navi Mumbai 410210, India
[d] Department of Radiation Oncology, University Medical Centre Utrecht, 3584 CX Utrecht, the Netherlands
[e] Department of Clinical Oncology, Leiden University Medical Center, 2333 ZA Leiden, the Netherlands
[f] Department of Radiation Oncology, Comprehensive Cancer Center, Medical University of Vienna, 1090 Vienna, Austria

Corresponding Authors: M. P. Email: marta.pelizzola@rm.dk

1. Factor analysis

Exploratory Factor Analysis (EFA) was used in this manuscript to identify clusters of symptoms. The Kaiser-Meyer-Olkin (KMO) test [Kaiser, 1970] and Bartlett's test [Bartlett, 1950] were used to assess the suitability of the data to EFA. The number of factors is chosen heuristically with the Kaiser rule, i.e. all factors with eigenvalue greater than 1 are included in the results [Kaiser, 1960]. Let N be the number of symptoms and $S = (S_1, \cdots, S_N)$ the observed symptoms and $\mu = (\mu_1, \cdots, \mu_N)$ the mean vector of S. Then the FA model is:

$$S = \mu + \ell \cdot F + \epsilon \tag{1}$$

where ℓ is the matrix of factor loadings (weights of the symptoms on the clusters in this framework), F is the matrix representing the factors and ϵ is the error term. Here, this model was applied with oblimin rotation on both EORTC and CTCAE data with symptoms and follow-ups as observations. Other methods such as hierarchical clustering [Murtagh and Contreras, 2012] have also been used to identify symptoms clusters.

Hierarchical clustering as other cluster analysis approaches reduces the dimensionality of the data and estimates clusters of similar variables by providing a grouping of the most similar symptoms. On the other hand, EFA estimates latent factors representing unobserved concepts explaining the joint variation of the observed variables. This approach is beneficial in this context as it reduces the number of variables thus simplifying the understanding of the data and it explains correlations between the observed variables through the unobserved factors.

2. Results

For the EORTC data, the KMO test returns a measure of sampling adequacy (MSA) of 0.89 for the entire data set and the Bartlett's test returns a p-value of $p < 0.001$, showing that this data is suited for factor analysis. The KMO test for CTCAE data returns MSA = 0.93 and the p-value for the Bartlett's test is $p < 0.001$, thus also the CTCAE data set is suited for EFA.

3. EMBRACE Collaborative Group

Department of Radiotherapy, Gustave-Roussy, Villejuif, France: Christine Haie-Meder, Isabelle Dumas, Cyrus Chargari.

Department of Oncology, Aarhus University Hospital, Denmark: Jacob C Lindegaard, Kari Tanderup, Lars Fokdal.

Department of Radiation Oncology, Comprehensive Cancer Center, Medical University of Vienna/General Hospital of Vienna, Vienna, Austria: Christian Kirisits, Alina Sturdza, Maximilian P Schmid.

Department of Radiation Oncology, University Medical Centre Utrecht, The Netherlands: Ina M. Jürgenliemk-Schulz, Astrid De Leeuw.

Department of Radiation Oncology, Tata Memorial Hospital, Mumbai, India: Umesh Mahantshetty, Jamema Swamidas, Supriya Chopra.

Department of Radiotherapy, Institute of Oncology Ljubljana, Slovenia: Barbara Segedin, Robert Hudej.

Department of Oncology, The Norwegian Radium Hospital, Oslo University Hospital, Oslo, Norway: Kjersti Bruheim, Taran Paulsen Hellebust.

Cancer Centre, Mount Vernon Hospital, London, UK: Peter Hoskin, Gerry Lowe.

Department of Radiotherapy and Oncology, Postgraduate Institute of Medical Education and Research, Chandigarh, India: Bhavana Rai, Arun S. Oinam.

Department of Oncology, Cross Cancer Institute and University of Alberta, Edmonton, Canada: Fleur Huang, Ericka Wiebe, Geetha Menon.

Leeds Cancer Centre, St James's University Hospital, Leeds, UK: Rachel Cooper.

Department of Radiotherapy, Arnhem, The Netherlands: Elzbieta Van Der Steen Banasik.

Department of Radiation Oncology, UZ Leuven, Belgium: Erik Van Limbergen. Clinic of Oncology and Women's Clinic, St. Olavs Hospital, Trondheim, Norway: Marit Sundset.

Department of Radiation Oncology Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands: Bradley Pieters.

Oncology Centre, Cambridge University Hospitals NHS Foundation Trust, Addenbrooke's Hospital, Cambridge, UK: Li Tee Tan.

Department of Radiation Oncology, Hospital of Navarra, Pamplona, Spain: Elena Villafranca.

Department of Radiation Oncology, Leiden University Medical Center, The Netherlands: Remi A. Nout.

Maastricht Radiation Oncology (MAASTRO) clinic, Maastricht, The Netherlands: Ludy C.H.W. Lutgens.

Department of Oncoradiology, University of Kaposvár, Health-sciences Center, Kaposvar, Hungary: Janaki Hadjiev.

Department of Radiation Oncology, British Columbia Cancer Agency, British Columbia, Canada: Francois Bachand.

Department of Radiation Oncology, Medical College of Wisconsin, Milwaukee, USA: Beth Erickson.

Department of Radiation Oncology, University of Iowa, Iowa, USA: Geraldine Jacobson.

Department of Obstetrics and Gynecology, Kuopio University Hospital, Finland: Maarit Anttila.

References

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Statistical Psychology*, 3:77–85.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1):141–151.

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, 35:401–415.

Murtagh, F. and Contreras, P. (2012). Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2:86–97.