European Journal of
Nuclear
Medicine

Editorial

Factor analysis revisited

David Barber and Anne Martel

Department of Medical Physics, Royal Hollamshire Hospital, Sheffield S10 2JF, UK

Scientific papers on the use of factor analysis (FA) in nuclear medicine have been finding their way into print for nearly 20 years. This issue contains two excellent examples. Much though not all of the work on factor analysis has been directed towards the analysis of dynamic studies (see the paper by Helal et al. in this issue), and it is this use of FA that we wish to discuss here, although other uses have also been published (see the paper by Szabo et al. in this issue). The drive behind much of this work has been the desire to do better than region-of-interest (ROI) analysis, the inadequacies of which are well known (or should be). Drawing regions has always been very operator dependent, and instances of variation between different operators are well documented. Cross-talk between regions often occurs, and the true amount of background to subtract is never really known. ROI analysis suffers from problems of both accuracy and reliability. FA appeared to offer at least partial solutions to some of these problems. However, data analysis techniques based on FA have still not found generally widespread use, although a few groups, such as that at Villejuif, have a long and distinguished record in the use of FA.

Why should this be so? Part of the reason must be the perceived mathematical difficulty of FA. This is not really fair. Matrix algebra, which forms the mathematical foundation of FA, is a standard tool in biophysics. However, it is true that there can be conceptual problems, such as trying to visualize oblique rotations of M-dimensional subspaces within an N-dimensional study space, which may be difficult for the novice user to cope with. We would guess, though we have no data to support this hypothesis, that groups which actively pursue FA have as a member an active subspace enthusiast. In spite of the apparent mathematical difficulties, the basic concept behind FA is simple. A radionuclide dynamic study often appears more complex than it really is. It turns out that the volume of data in a study can often be reduced by a factor of tenfold or better. The study can be compacted to a small set of factors (both images and curves), and these can be used at any time to reconstruct the original study without significant loss of information, although noise is reduced.

Unfortunately, the reduced data may not be in a form that can easily be interpreted. For example, the most powerful form of FA, in terms of the efficiency with which data can be reduced in volume, is principal components analysis (PCA). The factors produced by PCA cannot correspond directly with physiological factors in the data, so difficulty is experienced in interpreting them. Fairly early on it was recognized that we needed to get closer to the physiology than is allowed by PCA. It was realized that if the data could be described by a limited number of true physiological factors, then these factors would lie in the space defined by the principal components of the study. All that was needed was to find the position of these factors in this space. A substantial portion of the last 10 years of research into has been, at some risk of simplification, a search for a method of finding these factors. Since an infinite number of solutions are possible, constraints had to be used in order to find the true factors. The earliest constraint used was one which recognized that physically meaningful factors cannot have elements of negative value. This was the positivity constraint. Other constraints which have been proposed for calculating physiological factors include simple structure, clustering methods, spatial constraints and model fitting. For success each method requires certain assumptions about the structure of the data to be true. Unfortunately, examples can easily be found for each of these methods where the assumptions are not met. PCA produced factors of high efficiency but complete obscurity; these were not very user friendly. The positivity constraint produced solutions which were closer to the true physiological factors than PCA and these were a lot more friendly, or at least more understandable. However, the curves still did not represent the true factors, although their extraction was largely operator independent. Subsequent work has not produced any spectacular improvements on this. In our own laboratory we have investigated how well some constraints other than positivity can extract the true factors. Using conditions optimal for both FA and ROI analysis we have shown that FA can generate more accurate curves than ROI analysis. However, the gains are relatively small.

Where does this leave us? Can factor analysis improve the quality of dynamic study analysis? One of the negative aspects of FA has been the relatively long computing time required to compute the factors. Until people are convinced that FA can produce enough improvement in diagnostic performance to justify the extra

processing time, its widespread use will remain limited. However, at some point in the near future (it may be here now) computing the factors will become so rapid that the cost of obtaining them will be negligible. While we would argue that theoretically only limited improvements in data accuracy are possible using FA, there are other aspects of data quality which also require attention, such as reproducibility of results between different operators. Conditions are rarely perfect for ROI analysis. FA could possibly have a role in improving reproducibility. To take a simple but not trivial example, by determining approximate factors, it might be possible to use them as an aid in the drawing of reliable regions for a conventional ROI analysis! An original aim of FA was to take the human out of the analysis because he or she was known to be unreliable. Part of this unreliability was caused by the fact that he or she was given insufficient data to be able to perform optimally. One of the new aims of FA should be to find ways to put the human back into the analysis, but with sufficient information to make his or her contribution to the analysis much more reliable.

References

Helal BO, Frouin F, Schaison G, Leguillouzic D, Pueyo ME, Lebtahi R, Archambaud F, Desgrez A, Bazin JP, Di Paola R (1992) Diagnosis of malignancy in thyroid nodules by factor analysis of spectral and dynamic structures: a simultaneous dual-isotope dynamic study with thallium-201 and iodine-131. Eur J Nucl Med 19:517–521

Szabo Z, Camargo EE, Sostre S, Shafique I, Sadzot B, Links JM, Dannals RF, Wagner HN, Jr (1992) Factor analysis of regional cerebral glucose metabolic rates in healthy men. Eur J Nucl Med 19:469–475