What is the optimal cohort size for a reproducibility study of dynamic contrast-enhanced MRI?

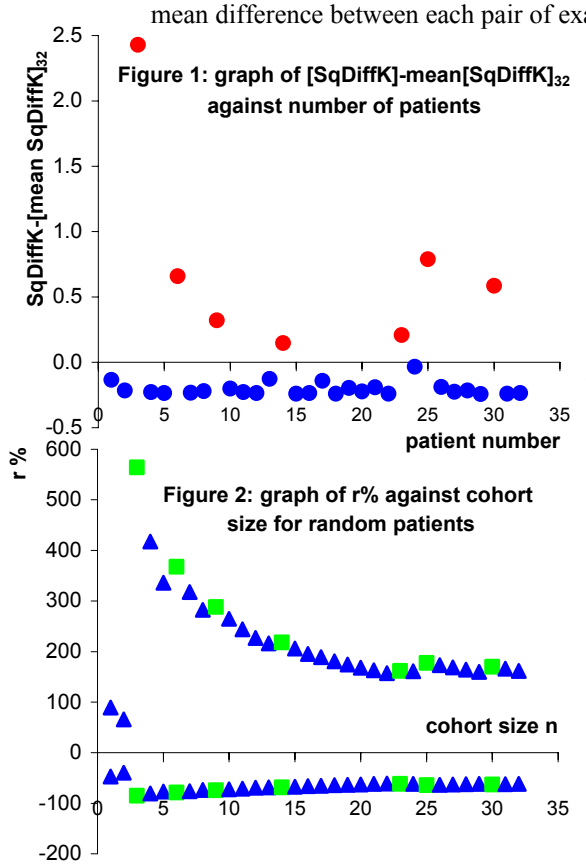
N. J. Taylor[1], J. A. d'Arcy[2], S. Walker-Samuel[2], D. J. Collins[2], B. Knowles[2], J. J. Stirling[1], T. Wallace[2], D-M. Koh[2], A. Tang[2], C. Lee[2], M. O. Leach[2], G. J. Rustin[3], A. R. Padhani[1]

[1]Paul Strickland Scanner Centre, Mount Vernon Hospital, Northwood, Middlesex HA6 2RN, United Kingdom, [2]Cancer Research UK Clinical Magnetic Resonance Research Group, Royal Marsden Hospital, Sutton, Surrey SM2 5PT, United Kingdom, [3]Academic Oncology, Mount Vernon Hospital, Northwood, Middlesex HA6 2RN, United Kingdom

The measurement of test-retest variability is now recommended as an integral part of study design of dynamic contrast enhanced MRI (DCE-MRI) studies that evaluate the effectiveness of antiangiogenesis and angiolytic drug trials [1]. This measurement of reproducibility enables investigators to define the level of change that would be statistically significant (for a single patient and larger cohorts). This information can then be used to identify a biologically active dose to take into efficacy studies and to assess the success of therapy. The purpose of this study was to assess the effect of increasing numbers of patients on the reproducibility of DCE-MRI studies in order to define the number of patients needed to adequately define test-retest variability.

Methods and results: 32 consecutive patients entered a two- centre anti-angiogenesis study; each imaging centre had more than 7 years of experience in performing DCE-MRI studies in support of clinical trials. Both centres used similar hardware (1.5T Siemens systems) and protocols: T_1weighted dynamic scans (three 8mm slices, TE 4.7ms, TR 11ms, α 30°) with a proton density reference image (TR 20 or 31ms, α 3°, rest as T1W) and 0.1mmol/kg Gd-DTPA injected at 4ml/s. The quality assurance and quality control procedures were identical for the 2 centres. Regions of interest were drawn by 2 experienced observers working independently and the data were analysed MRIW software (ICR, London) using the pharmokinetic model of Tofts with the vascular input function described by Weinmann et al. [2,3]. Tumour data were acquired on two occasions pre-treatment and used to calculate transfer constant (K^{trans}) for the whole tumour ROI from 3 slices sampling the tumour.

Patient data were ordered by the date at which they entered into the study and the repeatability parameter r (expressed as a % of the mean), which represents the 95% CI for change in 1 patient, was calculated as consecutive patients were added in to the maximum of 32 patients[4]. Data were transformed by natural logarithm as the mean difference between each pair of examinations was proportional to their means (positive Kendall's tau test) [5] In order to identify K^{trans} outliers we calculated the square differences in *ln* K^{trans} (SqDiffK) for consecutive patients and subtracted the mean SqDiffK of all 32 patients. We defined an outlier as a patient for whom this value was positive Fig 1 is a graph of [SqDiffK]-mean[SqDiffK]$_{32}$ against patient inclusion number which also allows the identification of outlier patients (7 of 32, shown in red●). The effect of outliers on the repeatability parameter (r%) with an increasing reproducibility cohort size is shown in Fig 2 (outliers indicated by green ■).



n	change in r% between n
1 to 5	29.8
6 to 10	6.1
11 to 15	3.6
16 to 20	3.5
21 to 25	-1.5
26 to 30	0.4

Table 1: Changes in the lower limit of the 95% CI between cohort sizes

The repeatability range (r%) for all 32 patients was -61.8% to 162.0% but this changed to -31.1% to 45.1% when outliers were excluded (n=25). In order to identify the optimal number of patients where the lower limit of reproducibility became flat we inspected the effect of increasing cohort size on the change in the lower 95% CI of r% (Table 1). This table shows that the lower curve changed little once 20 patients were accrued. Readers should note that it is the lower 95% CI that is of interest in clinical trials of anti-angiogenesis drugs.

Discussion: The test-retest variability of DCE-MRI data is dependent on many physical and physiological factors, which we do not explore in this abstract. We recognise that variability may be improved by a number of measures including improving SNR of images, by 3D imaging and by individualisation of vascular input function etc. However, we have shown that the effect of outliers on r% reduces as the number of patients in the cohort increases. Thus, having too few patients in the reproducibility cohort (fewer than ~15) risks having an imprecise estimate of the variability, which is prone to the detrimental effects of outliers (in this patient study, around 1 in 4-5 are outliers) whereas there is little additional benefit of having more than 20 patients to define the optimal lower limit of reproducibility. Analyses such as these could be performed as studies accrue patients to provide feedback as to when sufficient patient numbers have been obtained

References

[1] Leach MO *et al., Br J Cancer* 2005;**92(9)**:1599-1610.
[2] Tofts, P.S. and Kermode, A.G. *Magn Reson Med.* 1991; **17**: 357
[3] Weinmann H.J., *et al., Physiol Chem Phys Med NMR.* 1984;16(2):167-72.
[4] Bland, J. and Altman, D. *Br. Med. J.* 1996; **313**: 744
[5] Bland, J. and Altman, D. *Br. Med. J.* 1996; **313**: 106