Swiss sites was available for QC. According to a detailed study protocol, each site was requested to provide a range of various image reconstructions from multiple acquisitions of two differing phantoms. The software was applied in a single run to all datasets, producing an individual quality report for each dataset and a database making all results available for further evaluation. **Results**: The automated QC correctly classified the diverse datasets according to acquisition protocol, reconstruction protocol and type of phantom used. Corresponding protocol adherence verification and quantitative evaluations were performed. The system rejected 7% of the data sets because of serious violations of protocol. Another 57% of the datasets were categorized as of restricted use, because of a variety of minor violations of protocol or a mismatch between quantitative specifications and actual data. Only 36% of the data sets fully passed QC. The resulting database proved to be essential for further evaluations as published elsewhere. **Conclusion**: The results indisputably prove the need, the feasibility and the benefits of automated PET QC. More intelligent software is a prerequisite for rigorous PET QC of individual datasets, especially given the increasing diversity of PET imaging applications.

## EP-0048

### PET/CT scanner qualification in onco-haematological Clinical Trial: comparison between nationwide experience

*F. Bergesio[1], A. Biggi[2], M. Coronado[3], L. Ceriani[4], S. Chauvie[1]; [1]S. Croce e Carle Hospital, Medical Physics Unit, Cuneo, ITALY, [2]S. Croce e Carle Hospital, Nuclear Medicine Department, Cuneo, ITALY, [3]La Paz University Hospital, Nuclear Medicine Department, Madrid, SPAIN, [4]Oncology Institute of Southern Switzerland, Nuclear Medicine Department, Bellinzona, SWITZERLAND.*

**Purpose/Introduction:** The aim of this work was to compare the Clinical Trial Qualification (CTQ) adopted by the Italian Foundation on Lymphoma (FIL), the Grupo Espanol de Linfomas/Transplante Autologo de Medula Osea (GELTAMO) and the International Extranodal Lymphoma Study Group (IELSG). **Subjects & Methods:** Local personnel acquired Uniformity (UQP) and Image Quality (IQ) NEMA/IEC phantoms filled with 18F. The images were uploaded to a central server and analysed within Cuneo CoreLab. Background activity concentration (the average of activity concentration in a large homogenous ROI) in UQP (BACUQP) and in IQ (BACIQ) and sphere to background ratio (SBR) in the IQ phantom were compared to expected values. Recovery coefficient (RC) were calculated in the IQ phantom. Inter-scanner variability (ISV) of BACUQP, BACIQ and SBR was estimated as the 95% confidence level of difference between measured and expected values. Criteria for fulfilling the CTQ are a ISV less than 10% in BACUQP and RC coefficient within EANM limits. **Results:** 68/89 (76%) Italian PET/CT scanners fulfilled the CTQ. For qualified scanners the CTQ was reached at the first round in 35% of the cases, while in 31%, 16% and 18%, two, three or more than three iterations, were required, respectively. The ISV were 21.1%, 62.0% and 61.7% for BACUQP, BACIQ and SBR respectively. 25/26 (96%) Spanish PET/CT scanners fulfilled the CTQ. For qualified scanners the CTQ was reached at the first round in 24% of the cases, while in 40%, 16% and 20%, two, three or more than three iterations, were required, respectively. The ISV were 20.8%, 36.4% and 57.5% for BACUQP, BACIQ and SBR respectively. 22/29 (76%) worldwide PET/CT scanners fulfilled the CTQ. For qualified scanners the CTQ was reached at the first round in 59% of the cases, while in 27%, 0% and 14%, two, three or more than three iterations, were required, respectively. The ISV were 21.9%, 38.2% and 43.9% for BACUQP, BACIQ and SBR respectively. **Discussion/Conclusion:** The CTQ is a robust and reproducible procedure to verify inter-scanner calibration but has several limitations. Indeed, the BACIQ is at least two or three times larger than BACUQP demonstrating that the uniformity phantom, used to achieve CTQ, is more accurately prepared by local sites. SBR demonstrated a great variability (ISV from about 43.9% to 61.7%) because it accounts both for variability in phantom preparation and in reconstruction algorithm tuning.

## EP-0049

### Quantitative PET: Count Number Adaptations of Organ and Site specific Acquisition Protocols are a Key Determinant of comparable PET/CT Measurements

*G. A. Prenosil[1], T. K. Weitzel[1], M. Hentschel[1], T. Krause[1], A. Rominger[1], B. Klaeser[1,2]; [1]Department of Nuclear Medicine, Inselspital, Bern University Hospital, University of Bern, Switzerland, Bern, SWITZERLAND, [2]Department of Radiology and Nuclear Medicine, Winterthur Cantonal Hospital, Winterthur, SWITZERLAND.*

**Aim:** The ever-growing diversity and complexity of PET/CT systems make it increasingly difficult to define common acquisition protocols for quantifiable and comparable PET measurements in clinical routine and, above all, in multicenter clinical trials. We aimed to analyze the dependency of quantitative PET on technical variability, in order to improve acquisition and reconstruction protocols. The results formed the base for selected clinical acquisitions protocols, tailored to particular PET/CT systems. **Materials and Methods:** Image noise was examined in 422 PET/CT datasets from 18 PETCT systems participating in a Swiss multicenter phantom study. 215 of these measurements contained also hot spheres for recovery curve (RC) analysis. The study protocol combined different acquisition durations and low- and high-resolution image reconstructions with filtered back projection (FBP), ordered subset expectation maximization (OSEM), and vendor specific point spread function (PSF) based reconstruction. Data was analyzed with regard to exposure, defined as the product of background activity concentration and acquisition time. This produced results comparable in relation to the number of available decays per volume, independently from the actual activity concentration. Quality assurance (QA) limits for RCs were taken from guidelines issued by the Federal Office of Public Health in Switzerland. Additionally, RC shapes were quantitatively analyzed. **Results**: Passing the given QA limits depended highly on adequate exposure by keeping image noise levels low. Also, the minimal exposures required to fulfill the given QA limits differed between PET/CT systems, acquisition protocols and reconstruction algorithms. Despite higher image noise levels, quantifiability of FBP data was affected less by low count numbers than OSEM or PSF data. Insufficient count numbers led to faulty and incomparable image quantification,

and consequently also erratic RCs. From the gathered data, we were able to propose limits for minimal exposure for ten different PET/CT devices and for common clinical protocols, suitable for whole body and organ-specific high-resolution acquisitions. **Conclusions:** In order to reproducibly generate comparable and quantitative data, count numbers in PET/CT acquisitions must be specifically adapted to the clinically necessary image resolution, reconstruction method and the PET/CT system used. Therefore, optimal PET/CT acquisition times vary with the required exposure and with the expected activity uptake in target organs. This in turn argues against fixed acquisition times and mandates site and organ specific acquisitions protocols. Adapting exposure to a site's distinct technical factors appears to be a relevant element of PET/CT standardization in the context of multi-center trials.

## EP-0050

### Improving Standards Of Diagnostic Reporting In PET-CT: An Evaluation Of Factors Influencing Doctors' Performance On The National NHS England PET-CT Programme

*P. M. Ross[1], M. Frenz[1], M. F. Lambert[2], J. Armstrong[3], F. Gleeson[4], W. L. Wong[5]; [1]Birkbeck University of London, London, UNITED KINGDOM, [2]Public Health England, Newcastle upon Tyne, UNITED KINGDOM, [3]Alliance Medical Imaging [AML], London, UNITED KINGDOM, [4]Oxford University Hospitals Foundation Trust, Oxford, UNITED KINGDOM, [5]Mount Vernon hospital, Northwood, UNITED KINGDOM.*

**Aim:** 1**)** To identify and empirically measure type and number of reporting 'errors' in diagnostic PET-CT reports 2) To identify and evaluate likely reasons for variation in doctors' performance. **Materials and Methods:** The study examines two unique, merged datasets for 120 doctors: 1) Quantitative performance data from the National NHS England PET-CT Clinical Audit Programme pertaining to number and category of 'error' [discrepancy reports] found in audited reports between 2010 - 2016 2) Quantitative attitudinal data from a longitudinal survey to doctors reporting on the National NHS England PET-CT Programme 2013 - 2018 [National Programme]. Descriptive statistics were used to examine frequency and distribution of discrepancies in diagnostic reports. Regression techniques were applied to test for association between variation in doctors' performance with: activity volume, doctors' engagement [attitude toward audit], doctors' experience [measured as years reporting PET-CT, and as years reporting on the National Programme], 'rater' variability [defined as relative stringency or leniency with which an auditor rates diagnostic reports]. **Results:** Two distinct trends were observed in the performance data: 1) a positive skew in the frequency of errors, with some reporters increasingly associated with a higher percentage of discrepancy reports 2) a U-shaped pattern in the total percentage of discrepancies recorded across years which varied from 16% of all reports in 2010 to 3-5% in 2014 before increasing back to 13% in 2016. Proposition that positive skewness might be due to the higher absorptive capacity of some doctors which enabled them to learn more quickly from audit feedback. However, no reason for a U-shaped distribution of discrepancy reports was immediately evident. Robust regressions found positive associations with: 1) doctors' engagement in the audit process which explained up to 14% of variation in the model across 3 years 2) doctors' experience measured by years reporting on the National Programme which explained up to 8% of variation in model and 3) 'rater' variability which accounted for up to 46% of variation in the model between 2011- 2014 but only 6% in 2016. No support for an association with activity volume or experience measured as years reporting PET-CT. **Conclusion:** THe U-shaped variation in doctors' performance was partially explained by the change of auditor and the introduction of reporters with less experience of the National Programme. Regular auditor feedback and induction for new reporters irrespective of years of experience outside the National Programme, could significantly reduce the number of discrepancy reports.

## EP-0051

### Multicentre PET Standardization for an Amyloid Image Repository

*E. Prieto[1], P. B. Aguilar[1], V. Morán[2], M. I. Morales[1], J. Arbizu[1], J. M. Martí-Climent[1]; [1]Clínica Universidad de Navarra, Pamplona, SPAIN, [2]Clínica Universidad de Navarra, Madrid, SPAIN.*

From January 2017 until February 2018, a series of 12 phantom acquisitions were collected from different PET scanners in Spain, in order to check uniformity before starting a multicentre study for the amyloid neuroimaging repository of the PET-ADDs Consortium. **Objective:** To report the findings of a multicentre PET/CT phantom study for scanner standardization. **Methods:** Sites were required to complete a form, to acquire a PET/CT image of a Jaszczak phantom with hot spheres (contrast 10:1, diameters from 10 to 31 mm) following a detailed procedure and to submit the images for centralized evaluation. Average SUV in a spherical volume of interest in the background area of the phantom was used to determine SUV bias. A 10% deviation was stablished as acceptance criterion. SUV recovery coefficients (RC) for the six spheres were measured on 50% isocontour VOIS. Mean (SUV50) and maximum (SUVmax) values were recorded for each VOI. For a preliminary evaluation, RC values were compared to the EARL programme acceptance criteria, interpolating EARL RC for the Jaszczak diameters. **Results:** Datasets from 9 different scanner models were collected (42% from GE, 8% from Philips and 50% from Siemens). Five out of the 12 sites (42%) were out of the SUV bias tolerance of ± 10%. It is noteworthy that a deviation of 100% was found in one of the sites. These centres were asked to repeat the phantom scan and/or to take corrective actions (recalibration of the PET system). Regarding RC values, none of the SUV RC values of any centre was below the EARL acceptance criteria. However, only for one site all the RC values fell within the EARL limits and the remaining tomographs had 1 to 6 RC values above the tolerance. For the smallest sphere, observed RC values for SUVmax ranged from 0.42 to 0.81 (EARL acceptance 0.34-0.57) and RC SUV50 ranged from 0.28 to 0.55 (EARL acceptance 0.27-0.43). Centres were not asked for additional reconstructions and the tolerance limits are been reconsidered with available data for prospective site inclusion. **Conclusion:** This study has enabled the identification of SUV calibration errors in 42% cen-