18F-FDG PET/MRI for staging and interim response assessment in pediatric and adolescent Hodgkin lymphoma: a prospective study with 18F-FDG PET/CT as reference standard

Short running title
PET/MRI in pediatric Hodgkin lymphoma

Authors:
Verhagen, Martijn V, MD 1,2 m.verhagen@umcg.nl
Menezes, Leon J FRCP, FRCP 3,4 leon.menezes@nhs.net
Neriman, Deena, MB BS, FRCR 3 deena.neriman@nhs.net
Watson, Tom A, FRCR 2 tom.watson@gosh.nhs.uk
Punwani, Shonit, MRCP, FRCR, PhD 1 shonit.punwani@nhs.net
Taylor, Stuart A, MRCP, FRCR, MD 1, 5 stuart.taylor1@nhs.net
Shankar, Ananth, MRCP, MD, FRCPCH 5, 6 ananth.shankar@nhs.net
Daw, Stephen, FRCP 6 stephendaw@nhs.net
Humphries, Paul D, MRCP, FRCR 1,2* paul.humphries5@nhs.net

*Corresponding author

1 Department of Radiology
University College London Hospital, 235 Euston Road, London, UK NW1 2BU
2 Department of Radiology
Great Ormond St, London, UK WC1N 3JH

PET/MRI in pediatric Hodgkin lymphoma
3 UCL Institute of Nuclear Medicine
University College London Hospital, 235 Euston Road, London, UK NW1 2BU

4 NIHR University College London Hospitals Biomedical Research Centre

5 Centre for Medical Imaging, CBH
43-45 Foley Street, London, UK W1W 7TS

6 Department of pediatrics
University College London Hospital, 235 Euston Road, London, UK NW1 2BU

Disclaimer
Funding was secured from Great Ormond Street Hospital Childrens' Charity.

First author
Verhagen, Martijn V, MD. Pediatric radiologist

m.verhagen@umcg.nl
Phone: +31(0)625649708
Fax: +31 (050) 361 37 55

Address
University Medical Center Groningen (UMCG)
Medical Imaging Center, EB44
Hanzeplein 1, 9713GZ Groningen
The Netherlands

PET/MRI in pediatric Hodgkin lymphoma
Corresponding author

Humphries, Paul D, MRCP, FRCR 1,2* paul.humphries5@nhs.net

Address correspondence to:

Humphries, Paul D. Department of Radiology, University College London Hospital, 235 Euston Road, London, UK NW1 2BU
Tel: +44 (020) 7405 9200
Fax: +44 (020) 7813 8150

Statement

Funding was secured from Great Ormond Street Hospital Childrens’ Charity.
Title: 18F-FDG PET/MRI for staging and interim response assessment in pediatric and adolescent Hodgkin lymphoma: a prospective study with 18F-FDG PET/CT as reference standard

ABSTRACT

Rationale:
Treatment regimens for pediatric Hodgkin’s lymphoma (HL) depend on accurate staging and treatment response assessment, based on accurate disease distribution and metabolic activity depiction. With the aim of radiation dose reduction, we compared the diagnostic performance of 18F-FDG PET/MR to a 18F-FDG PET/CT reference standard for staging and response assessment.

Methods:
Twenty-four patients (mean age 15.4 years, range 8-19.5 years) with histologically proven HL were prospectively and consecutively recruited in 2015 and 2016, undergoing both 18F-FDG PET/CT and 18F-FDG PET/MRI at initial staging (N = 24) and at response assessment (N = 21). Diagnostic accuracy of 18F-FDG PET/MRI for both nodal and extra-nodal disease was compared to 18F-FDG PET/CT, which was considered as the reference standard. Discrepancies were retrospectively classified as perceptual or technical errors and 18F-FDG PET/MRI and 18F-FDG PET/CT were corrected by removing perceptual error. Agreement with Ann-Arbor staging and Deauville grading was also assessed.
Results:
For nodal and extranodal sites combined, corrected staging $^{18}\text{F-FDG}$ PET/MRI sensitivity was 100% (95% confidence interval (CI) 96.7%-100%), specificity 99.5% (95%CI 98.3%-99.9%). Corrected response assessment $^{18}\text{F-FDG}$ PET/MRI sensitivity was 83.3% (95%CI 36.5%-99.1%), specificity 100% (95%CI 99.2%-100%). Modified Ann-Arbor staging agreement between $^{18}\text{F-FDG}$ PET/CT and $^{18}\text{F-FDG}$ PET/MRI was perfect ($k = 1.0, p = 0.000$). Deauville grading agreement between $^{18}\text{F-FDG}$ PET/MRI and $^{18}\text{F-FDG}$ PET/CT was excellent ($k = 0.835, p = 0.000$).

Conclusion:
$^{18}\text{F-FDG}$ PET/MRI is a promising alternative to $^{18}\text{F-FDG}$ PET/CT for staging and response assessment in children with Hodgkin lymphoma.

Key words:
Hodgkin lymphoma; PET-MRI; PET-CT; pediatric; staging
INTRODUCTION

Hodgkin lymphoma (HL) is the most common childhood lymphoma, and one of the most common pediatric and adolescent cancers (1). Treatment outcomes are critically dependent on accurate imaging defined staging and thereafter, early treatment response assessments.

Staging consists of detection of anatomical locations of disease, using either CT or MRI, combined with metabolic assessment (2). This metabolic assessment consists of using 18F-2-deoxy-2-fluoro-D-glucose Positron Emission Tomography (18F-FDG-PET) to detect increased glucose metabolism of lymphatic and extra-lymphatic tissue involved by HL (3). Imaging is repeated after two (for classical HL) or three (for lymphocyte predominant HL) cycles of chemotherapy to determine early response assessment, which determines if post-treatment radiotherapy is necessary (3). The decision to consolidate with radiotherapy is taken if there is either insufficient reduction in size of disease or if there is persistent abnormal metabolic activity. It is well established that the addition of 18F-FDG-PET to standard cross-sectional imaging improves accuracy for both initial staging and treatment response assessment, so most patients undergo repeated multi-modality imaging.

Both CT and 18F-FDG PET impart a significant radiation dose and repeated studies result in a cumulative dose associated with secondary malignancies (4,5). Children and young adults are particularly vulnerable to the long term effects of radiation and according to the linear no threshold model of radiation, any reduction in radiation exposure for this group would be beneficial (6,7). With hybrid 18F-FDG PET/MRI in pediatric Hodgkin lymphoma
PET/MRI systems now available and able to potentially replace $^{18}$F-FDG PET/CT with a single $^{18}$F-FDG PET/MRI (8,9), diagnostic radiation exposure could potentially be reduced by up to 80% (5,8,10), whilst maintaining both anatomical and metabolic information.

We aimed to test the hypothesis that $^{18}$F-FDG PET/MRI is an alternative to $^{18}$F-FDG PET/CT, for both staging and early response assessment. This was done by prospectively assessing and comparing the diagnostic performance of $^{18}$F-FDG PET/MRI against standard $^{18}$F-FDG PET/CT for initial staging and early response assessment in pediatric and adolescent patients with HL.

MATERIALS AND METHODS

Patient Population

This single centre prospective study was approved by the institutional review board, and written informed consent was obtained from either participant and/or guardians in all patients.

Patients were consecutively included between February 2015 and June 2016. Inclusion criteria were; histologically confirmed Hodgkin lymphoma, age 0-20 years, inclusion into the EuroNet PHL-C1 trial, LP1 trial or successor Euronet trials including Euronet PHL C2. Exclusion criteria were; previous HL, prior chemotherapy or radiotherapy, pregnancy, breastfeeding and contra-indications to MRI.

Summary of the Study

As part of standard of care staging imaging at our institution, all patients underwent $^{18}$F-FDG PET/CT. The trial $^{18}$F-FDG PET/MRI was performed at the same PET/MRI in pediatric Hodgkin lymphoma
attendance as $^{18}$F-FDG PET/CT, using the same $^{18}$F-FDG injection, with the patient transferring to the $^{18}$F-FDG PET/MRI suite immediately after completion of the $^{18}$F-FDG PET/CT examination.

Following two (EuroNet PHL-C1) or three (EuroNet LP1) cycles of chemotherapy, disease was reassessed with $^{18}$F-FDG PET/CT and the trial $^{18}$F-FDG PET/MRI again performed immediately after the $^{18}$F-FDG PET/CT. $^{18}$F-FDG PET/CT was the standard of reference against which $^{18}$F-FDG PET/MRI was compared.

$^{18}$F-FDG PET/CT Protocol

$^{18}$F-FDG PET/CT data was acquired using GE Discovery PET/CT in-line systems (General Electric Healthcare, Milwaukee, Wisconsin, USA) without time-of-flight (TOF) technology. Patients fasted for 6 hours, and hyperglycaemia (>10 mmol/L) was excluded. A weight-adjusted dose of $^{18}$F-FDG was injected 60 minutes before imaging. Patients were scanned from skull vertex to mid-thigh level, at 3 minutes per bed position. CT was low dose for attenuation correction, acquisition parameters were adjusted according to patient weight. No intra-venous iodinated contrast was administered. Combined axial emission images of $^{18}$F-FDG PET and CT were reconstructed to a 128x128 resolution image, with 2.5mm slice thickness.

$^{18}$F-FDG PET/MRI Protocol

$^{18}$F-FDG PET/MRI sequences were acquired on a hybrid 3 Tesla MR without TOF (Siemens Healthineers Biograph mMR, Erlangen, Germany). All sequences were acquired from skull vertex to mid-thigh level, at 5 minutes per bed position.

$^{18}$F-FDG PET/MRI comprised of axial T2-HASTE (half-Fourier acquisition single-shot turbo spin-echo), axial and coronal TIRM (turbo inversion recovery
magnitude) sequences, and axial post Gadolinium T1 Dixon fat suppression sequences (TABLE 1). Four tissue class (soft tissue, fat, lung, air) attenuation correction maps were calculated from the two-point Dixon sequences. Combined axial emission images of $^{18}$F-FDG PET and T2-HASTE sequences were reconstructed to a 128x128 resolution image, with 5 mm slice thickness.

**Analysis of Images**

All studies were read sequentially under trial conditions using Osirix workstation software (Osirix MD, Geneva, Switzerland) after all scans were performed. Readers were blinded to patient data, clinical data, and other modalities including the reference standard.

$^{18}$F-FDG PET/CT scans were read by a nuclear medicine specialist (DN, 10 years dedicated nuclear medicine experience). $^{18}$F-FDG PET/MRI sequences were evaluated by a nuclear medicine specialist and pediatric radiologist in consensus (LM and PDH, both 15 years of experience).

Derivation of enhanced reference standard and correction for perceptual errors

Discrepancies between $^{18}$F-FDG PET/MRI and the $^{18}$F-FDG PET/CT reference standard were reviewed by an independent non-blinded pediatric radiologist (MV, 5 years dedicated pediatric radiology experience), using similar methodology to that of Latifoltojar et al (11). Minor discrepancies resulting from disease differently located at site boundaries were not labelled as discrepant. All other discrepancies were reviewed in consensus between MV and PH and assigned to error type ‘perceptual’ or ‘technical’. Perceptual errors entailed retrospectively visible disease. A technical
error was noted for discrepancies unrelated to reader detection, such as difference in measured metabolic activity between $^{18}$F-FDG PET/MRI and $^{18}$F-FDG PET/CT.

Technical discrepancies were always considered in favour of the $^{18}$F-FDG PET/CT reference standard. By comparing datasets corrected for perceptual error we aimed to remove the human perception bias and better compare technical performance.

**Staging Definitions**

Staging was performed according to the Cotswolds modified Ann Arbor classification (12,13). Nineteen nodal sites were assessed: cervical (left and right), anterior mediastinum, paratracheal, lung hilum (left and right), diaphragm, axilla (left and right), hepatic hilum, splenic hilum, spleen, coeliac trunk, para-aortic, mesenteric, iliac (left and right), inguinal (left and right). Ten extra-nodal sites were also registered: Lung, chest wall, kidneys, bone marrow, pleural, pericardium, bowel, stomach, liver and pancreas.

Nodal and extra-nodal disease involvement was defined as a focal area of FDG$^{18}$ with a SUVmax (measured using a 1 cm$^2$ circular region of interest) uptake above mediastinal blood pool SUVmax (14) or FDG uptake greater than the surrounding background in a location incompatible with normal physiological activity (15,16). Nodal volumes were derived from measurements in 3 orthogonal dimensions (volume = $1/6 \pi (a \times b \times c)$).

Any focal lung parenchymal $^{18}$F-FDG avid focus above mediastinal blood pool SUVmax was considered involved. As per EuroNet criteria, focal lung parenchymal
involvement was also considered involved in case of a focal consolidation, solitary nodule of greater than 1 cm or more than three sub-centimeter nodules.

Extra-nodal extension of disease was registered for contiguous extension of tissue beyond a nodal mass into adjacent structures.

Bone marrow involvement at staging was concluded if focal or multifocal FDG uptake was above that of the liver (17), at response assessment bone marrow involvement was defined as focal or multifocal higher uptake compared to normal marrow but reduced compared with baseline (with diffuse changes from chemotherapy allowed) (18). Bone marrow lesions on MRI were only considered as bone marrow involvement when combined with increased FDG uptake.

For response assessment each site was re-evaluated and residual metabolic activity classified using the 5-point Deauville scale (19).

**Statistical Analysis**

Detection rates (sensitivity and specificity, with a 95% confidence interval) and area under the curve (AUC) were calculated for nodal and extra-nodal sites combined. $^{18}$F-FDG PET/MRI was subsequently compared to $^{18}$F-FDG PET/CT as reference standard; with and without correction for perceptual error.

The kappa statistic was determined to test agreement between modalities. This was classified as poor for $\kappa < 0.00$; slight for $\kappa 0.00–0.20$; fair for $\kappa 0.21–0.40$; moderate for $\kappa 0.41–0.60$; good for $\kappa 0.61–0.80$; and excellent for $\kappa 0.81–1.00$ (20).

Individual nodal volume and individual SUVmax of nodal and extra-nodal sites were compared between $^{18}$F-FDG PET/MRI and $^{18}$F-FDG PET/CT at staging and response assessment. Correlation was calculated using Spearman’s correlation test.
Statistics were performed using SPSS for Windows (release 26, IBM, New York, United States of America). The level of significance was set at $\alpha < 0.05$.

RESULTS

Twenty-six patients were recruited (male:female ratio 14:12, median age 16, range 8-19 years). After exclusion, 24 patients were enrolled for staging, and 21 remained also for response assessment. Twenty-two patients had histologically confirmed classic HL, two had nodular lymphocyte predominant HL. The study flowchart is presented in FIGURE 1.

The median time between $^{18}$F-FDG injection and $^{18}$F-FDG PET/CT was 68 minutes (interquartile range (IQR) 16). Median interval time between $^{18}$F-FDG PET/CT and $^{18}$F-FDG PET/MRI was 95 minutes (IQR 35.5).

Staging

$^{18}$F-FDG PET/CT detected 141 disease positive sites (136/456 nodal, 5/240 extranodal: 696 sites total) (TABLE 2). The modified Ann Arbor stage distribution was: stage 1, 1 patient (1/24: 4.2%); stage 2, 13 patients (13/24: 54.2%); stage 3, 6 patients (6/24: 25%), stage 4, 4 patients (4/24: 16.7%).

$^{18}$F-FDG PET/MRI detected 135 out of 141 true positive sites, 551 true negative sites, 4 false positive sites, and 6 false negatives (TABLE 2). Six out of 6 false negative sites were because of perceptual error. Three out of 4 false positive sites were because of technical errors, these false positive sites were small in volume (0.02-1.3ml) but $^{18}$F-FDG avid, whereas they were not $^{18}$F-FDG avid on the $^{18}$F-FDG PET/MRI in pediatric Hodgkin lymphoma.
PET/CT reference standard. The remaining single false positive site was because of perceptual error (TABLE 2).

Uncorrected and corrected sensitivity and specificity are given in TABLE 3. $^{18}$F-FDG PET/MRI intermodality agreement for disease sites compared to $^{18}$F-FDG PET/CT detection was excellent for both uncorrected and corrected data (TABLE 3).

There were no discrepancies between $^{18}$F-FDG PET/MRI and $^{18}$F-FDG PET/CT for modified Ann Arbor staging (TABLE 4). Figure 2 demonstrates an involved nodal site at staging, which was concordant between $^{18}$F-FDG PET/CT and $^{18}$F-FDG PET/MRI.

**Response Assessment**

At response assessment, $^{18}$F-FDG PET/CT demonstrated 6 incomplete metabolic response sites (4/399 nodal and 2/210 extranodal: 609 sites total) in 3 out of 21 patients (TABLE 2). Deauville grade distribution was: grade 2, 11 patients (11/21: 52.4%); grade 3, 7 patients (7/21: 33.3%); grade 4, 3 patients (3/21: 14.3%). Figure 3 demonstrates $^{18}$F-FDG PET/MR of nodal site involvement at staging and response assessment.

$^{18}$F-FDG PET/MRI correctly detected 5/6 incomplete metabolic response sites, 1 incomplete metabolic response site was not detected due to a technical error (Deauville 3 on $^{18}$F-FDG PET/MRI compared to Deauville 4 on $^{18}$F-FDG PET/CT). There were no perceptual errors. Uncorrected and corrected sensitivity and specificity are given in TABLE 3.
Intermodality agreement for disease sites between $^{18}$F-FDG PET/MRI and $^{18}$F-FDG PET/CT detection was excellent (TABLE 3). $^{18}$F-FDG PET/MRI response assessment according to Deauville was excellent (TABLE 4).

**Extra-Nodal Disease**

At staging 5/141 involved sites were extranodal (3 bone marrow sites, and 2 lung sites). Regarding the lung sites, one patient had one lung nodule measuring 12 and 14 mm on T2-HASTE and CT, respectively. The other patient had multiple lung nodules: 12 nodules were detected on T2-HASTE measuring up to 14 mm, and 41 nodules measuring up to 18 mm on CT.

At response assessment 2 incomplete metabolic response sites were extranodal (one bone marrow and one lung site). Further statistical analyses was omitted due to the low number of extranodal sites.

**Volume**

At staging, correlation between individual nodal disease sites volumes at $^{18}$F-FDG PET/MRI and $^{18}$F-FDG PET/CT was excellent ($r = 0.817, p = 0.000$). Volumes are given in TABLE 5.

At response assessment, correlation was good ($r = 0.728, p = 0.000$). Volumes and volume reduction percentages at response assessment are given in TABLE 5.

**SUVmax**

At staging, there was good correlation ($r = 0.732, p = 0.000$) between SUVmax from $^{18}$F-FDG PET/MRI and $^{18}$F-FDG PET/CT, SUVmax values are given in TABLE 5.
Correlation at response assessment was moderate ($r = 0.543$, $p = 0.000$), SUVmax values are given in TABLE 5.

**DISCUSSION**

This study prospectively compared $^{18}$F-FDG PET/MRI to $^{18}$F-FDG PET/CT for both staging and early post-chemotherapy response reassessment in a cohort of children and adolescents with Hodgkin Lymphoma.

By comparing the diagnostic performance of these modalities, we wanted to test the hypothesis that $^{18}$F-FDG PET/MRI is an alternative to $^{18}$F-FDG PET/CT in pediatric HL, with the aim of reducing the cumulative radiation dose that these children receive throughout their illness. To correct for human perception bias, we present data corrected and uncorrected for perceptual error. Technical error due to differences in technique was not corrected for.

When corrected for perceptual error our results showed perfect agreement for modified Ann Arbor staging between $^{18}$F-FDG PET/MRI and $^{18}$F-FDG PET/CT, and excellent response assessment agreement according to the Deauville scale. The notion that $^{18}$F-FDG PET/MRI may replace $^{18}$F-FDG PET/CT is further supported by good to excellent intermodality correlation of SUVmax and nodal sizes at staging, and moderate to good correlation at response assessment.

In adults, excellent to perfect staging agreement ($k = 0.979$-$1.0$) between $^{18}$F-FDG PET/CT and $^{18}$F-FDG PET/MRI for staging of mixed groups of Hodgkin and non-Hodgkin lymphoma has been reported ($13,21,22$). However, given the differences in body habitus, physiology (e.g. brown fat), the challenges of prolonged MRI protocols in children, and possibly different behaviour of lymphomas, adult PET/MRI in pediatric Hodgkin lymphoma
studies comparing $^{18}$F-FDG PET/MRI with $^{18}$F-FDG PET/CT cannot be extrapolated
to children and adolescents (4,11,23). Those available pediatric studies comparing
$^{18}$F-FDG PET/MRI and $^{18}$F-FDG PET/CT have so far consisted of a mix of both
Hodgkin and non-Hodgkin lymphoma, and studied either staging or response
assessment (4,24). Our study aimed to improve on the pediatric literature by
prospectively including only pediatric HL and studying both staging and early
response assessment.

Because of a longer study duration, a smaller bore diameter, and loud noises,
undergoing $^{18}$F-FDG PET/MRI instead of $^{18}$F-FDG PET/CT may be challenging for
some children (and adults). Preparation using virtual reality glasses, a mock MRI and
play specialists may help children to be comfortable in an MRI (25). However,
diagnostic performance must have priority over reduction of radiation dose. For those
children unable to lie still in the MRI, either anesthesia or $^{18}$F-FDG PET/CT instead of
$^{18}$F-FDG PET/MRI may be necessary.

In addition to reduced radiation dose, replacing $^{18}$F-FDG PET/MRI with $^{18}$F-
FDG PET/CT may have another benefit. MRI sequences have intrinsic superior soft
tissue contrast compared to CT, and this may allow for better delineation of lymph
nodes (e.g. hilar lymph nodes) and focal solid organ lesion. However, although lung
nodules can be seen on MRI, CT has superior air to tissue contrast, and a diagnostic
chest CT is advised at staging of lung nodules (26). Consequently, our current
practise consists of $^{18}$F-FDG PET/MRI at staging and response assessment, in
combination with a diagnostic non-contrast enhanced chest CT at staging in all
patients, and at response assessment in those children with lung involvement at staging.

A limitation of this study was the necessity to scan and then move the patients between the $^{18}$F-FDG PET/CT and $^{18}$F-FDG PET/MRI. This was again unavoidable where ethics allows only one $^{18}$F-FDG injection.

The effects of delay between injection and PET scanning are twofold. The first is reduction in $^{18}$F-FDG activity due to decay of the isotope, which may influence SUV measurements. The second is prolonged uptake time, which may increase the detection of disease locations due to washout in normal structures and higher activity in some disease sites due to slow tracer uptake (27). There are published examples of this phenomenon, for example false negative $^{18}$F-FDG PET/CT but true positive $^{18}$F-FDG PET/MRI adrenal lesions (13), which were thought be the result of prolonged uptake time.

Although we found no difference in staging between $^{18}$F-FDG PET/MRI and $^{18}$F-FDG PET/CT at diagnosis, there were site discrepancies which could have a potential clinical impact for children requiring radiotherapy. These discrepancies, regarded in our study as technical errors, may be due to the delayed $^{18}$F-FDG PET/MRI compared to $^{18}$F-FDG PET/CT. In our study, three (small volume: range 0.02-1.3ml) sites were considered false positive on $^{18}$F-FDG PET/MRI at staging due to higher SUVmax on $^{18}$F-FDG PET/MRI compared to $^{18}$F-FDG PET/CT. Although the reason for this cannot be verified, we postulate that this discrepancy may be related to the prolonged uptake time on the $^{18}$F-FDG PET/MRI. Conversely, differences in perfusion and washout may also account for the lower Deauville scale measurement of 2 sites on $^{18}$F-FDG PET/MRI during response assessment.
Lastly, there were only 6 sites at response assessment which did not show complete metabolic response. Although there was excellent agreement of $^{18}$F-FDG PET/MRI compared to $^{18}$F-FDG PET/CT at response assessment, these numbers are too small to draw definitive conclusions. Further studies to confirm these findings should include a larger population, which may be feasible especially in a multi-center collaboration.

CONCLUSION

In our cohort of patients, $^{18}$F-FDG PET/MRI showed no difference in overall staging compared to $^{18}$F-FDG PET/CT for staging of HL in children and adolescents, and there was excellent response assessment agreement. With the aim of reducing cumulative radiation dose, we suggest that pediatric/adolescent HL staging and response assessment may be performed using $^{18}$F-FDG PET/MRI instead of $^{18}$F-FDG PET/CT, wherever possible.
KEY POINTS

QUESTION:
- Can $^{18}$F-FDG PET/MRI replace $^{18}$F-FDG PET/CT for staging and response assessment of pediatric Hodgkin lymphoma?

PERTINENT FINDINGS:
- This prospective observational study demonstrates that $^{18}$F-FDG PET/MRI is a promising alternative to $^{18}$F-FDG PET/CT for staging and chemotherapy response assessment for pediatric Hodgkin lymphoma.

IMPLICATIONS FOR PATIENT CARE:
- By replacing $^{18}$F-FDG PET/CT with $^{18}$F-FDG PET/MRI, children with Hodgkin lymphoma will receive a lower cumulative radiation dose at staging and response assessment, whilst maintaining diagnostic accuracy.

DISCLOSURES

Funding was secured from Great Ormond Street Hospital Childrens’ Charity.

None of the authors have any potential financial conflict of interest related to this manuscript.
REFERENCES


18. Barrington SF, Kluge R. FDG PET for Therapy monitoring in Hodgkin and non-


## TABLE 1. MRI sequences

<table>
<thead>
<tr>
<th></th>
<th>TIRM</th>
<th>Dixon</th>
<th>T2-HASTE</th>
<th>3D VIBE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scanning plane</td>
<td>axial/cor</td>
<td>axial</td>
<td>axial</td>
<td>axial (liver and spleen)</td>
</tr>
<tr>
<td>Fat suppression</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Inversion time (ms)</td>
<td>220</td>
<td>N/A</td>
<td>N/A</td>
<td>180</td>
</tr>
<tr>
<td>Gadolineum-chelate</td>
<td>N/A</td>
<td>yes</td>
<td>N/A</td>
<td>0, 30, 60 sec</td>
</tr>
<tr>
<td>Echo Time (ms)</td>
<td>56</td>
<td>1.23, 2.46</td>
<td>92</td>
<td>1.91</td>
</tr>
<tr>
<td>Repetition Time (ms)</td>
<td>6870</td>
<td>4.1</td>
<td>1500</td>
<td>4.3</td>
</tr>
<tr>
<td>Flip Angle (degrees)</td>
<td>135</td>
<td>9</td>
<td>102</td>
<td>9</td>
</tr>
<tr>
<td>Section Thickness (mm)</td>
<td>7/4</td>
<td>3</td>
<td>5</td>
<td>3.5</td>
</tr>
<tr>
<td>Spacing (mm)</td>
<td>8.75/5.20</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>B-value s/mm²</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Echo train</td>
<td>11</td>
<td>1</td>
<td>256</td>
<td>1</td>
</tr>
<tr>
<td>Pixel space (mm)</td>
<td>1.76 x 1.76</td>
<td>0.78 x 0.78</td>
<td>1.02 x 1.02</td>
<td>1.19 x 1.19</td>
</tr>
<tr>
<td>Resolution (voxel)</td>
<td>256 x 176</td>
<td>640 x 500</td>
<td>448 x 336</td>
<td>320 x 260</td>
</tr>
</tbody>
</table>

Cor: coronal
HASTE: half-Fourier acquisition single-shot turbo spin-echo
mm: millimetre
ms: milliseconds
N/A: not applicable
TIRM: turbo inversion recovery magnitude
VIBE: volumetric interpolated breath-hold examination
TABLE 2. Technical and perceptual errors, nodal and extra-nodal sites combined

<table>
<thead>
<tr>
<th></th>
<th>Staging</th>
<th>Response assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>¹⁸F-FDG PET/CT(n)</td>
<td>¹⁸F-FDG PET/MRI(n)</td>
</tr>
<tr>
<td>TP</td>
<td>141</td>
<td>135</td>
</tr>
<tr>
<td>FP</td>
<td>N/A</td>
<td>4</td>
</tr>
<tr>
<td>- FP perceptual</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>- FP technical</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>TN</td>
<td>560</td>
<td>551</td>
</tr>
<tr>
<td>FN</td>
<td>N/A</td>
<td>6</td>
</tr>
<tr>
<td>- FN perceptual</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>- FN technical</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Total number of</td>
<td>696</td>
<td>696</td>
</tr>
</tbody>
</table>
sites

FN: false negative; FP: false positive; N/A: not applicable; n: number of patients; TN: true negative; TP: true positive
TABLE 3. ¹⁸F-FDG PET/MRI detection of involved sites

<table>
<thead>
<tr>
<th></th>
<th>Uncorrected for perceptual error</th>
<th>Corrected for perceptual error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td></td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>Staging</td>
<td>95.7% (135/141, 95%CI 90.6%-98.3%)</td>
<td>100% (141/141, 95%CI 96.7%-100%)</td>
</tr>
<tr>
<td></td>
<td>99.3% (551/555, 95%CI 98.0%-99.8%)</td>
<td>99.5% (552/555, 95%CI 98.3%-99.9%)</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.960 (p = 0.000)</td>
<td>0.987 (p = 0.000)</td>
</tr>
<tr>
<td>AUC</td>
<td>0.979 (p = 0.000)</td>
<td>0.997 (p = 0.000)</td>
</tr>
<tr>
<td></td>
<td>83.3% (5/6, 95%CI 36.5%-99.1%)</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>100% (603/603, 95%CI 99.2%-100%)</td>
<td>N/A</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.908 (p = 0.000)</td>
<td>N/A</td>
</tr>
<tr>
<td>AUC</td>
<td>0.917 (p = 0.000)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

AUC: area under the curve; 95%CI: 95% confidence interval; N/A: not applicable in absence of perceptual errors at response assessment
TABLE 4. Staging and early response assessment, agreement of $^{18}$F-FDG PET/MRI compared to $^{18}$F-FDG PET/CT

<table>
<thead>
<tr>
<th></th>
<th>$^{18}$F-FDG PET/MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Staging ($N = 24$)</strong></td>
<td></td>
</tr>
<tr>
<td>Concordant</td>
<td>24</td>
</tr>
<tr>
<td>Upstaged</td>
<td>0</td>
</tr>
<tr>
<td>Downstaged</td>
<td>0</td>
</tr>
<tr>
<td>Intermodality agreement (kappa)</td>
<td>1.0 ($p = 0.000$)</td>
</tr>
<tr>
<td><strong>Response ($N = 21$)</strong></td>
<td></td>
</tr>
<tr>
<td>Concordant</td>
<td>19</td>
</tr>
<tr>
<td>Higher Deauville</td>
<td>0</td>
</tr>
<tr>
<td>Lower Deauville</td>
<td>2*</td>
</tr>
<tr>
<td>Intermodality agreement (kappa)</td>
<td>0.835 ($p = 0.000$)</td>
</tr>
</tbody>
</table>

*One case because of Deauville 3 instead of 4, and one because of Deauville 2 instead of 3

N = number
TABLE 5. Volume and SUVmax per individual nodal sites

<table>
<thead>
<tr>
<th></th>
<th>(^{18}\text{F-FDG PET/MRI})</th>
<th>(^{18}\text{F-FDG PET/CT})</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Volume</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staging</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median, ml (IQR)</td>
<td>6.5 (19.2)</td>
<td>8.3 (26.7)</td>
</tr>
<tr>
<td>Response</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median, ml (IQR)</td>
<td>1.4 (2.9)</td>
<td>1.8 (5.7)</td>
</tr>
<tr>
<td>Volume reduction, %</td>
<td>86.90</td>
<td>80.35</td>
</tr>
<tr>
<td><strong>SUVmax, mg/ml</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staging</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median, mg/ml (IQR)</td>
<td>8.0 (4.8)</td>
<td>8.2 (4.8)</td>
</tr>
<tr>
<td>Response</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median, mg/ml (IQR)</td>
<td>1.2 (0.6)</td>
<td>1.7 (0.7)</td>
</tr>
</tbody>
</table>

IQR: Interquartile range; N/A: Not applicable; SD: Standard deviation;

SUVmax: maximum standard uptake value
FIGURE LEGENDS

FIGURE 1
Study Flowchart

Patients recruited
$N = 26$

Excluded $N = 2$
- No HL on pathology review ($N = 1$)
- No $^{18}$F-FDG PET/MRI ($N = 1$)

$^{18}$F-FDG PET/MRI and $^{18}$F-FDG PET/CT, $N = 24$

Treatment

Excluded for response assessment $N = 3$
- No $^{18}$F-FDG PET/MRI

$^{18}$F-FDG PET/MRI and $^{18}$F-FDG PET/CT, $N = 21$

Analysis 1
Concordance between $^{18}$F-FDG PET/MRI and $^{18}$F-FDG PET/CT
- nodal and extra-nodal sites at staging and response assessment
- Ann Arbor staging at staging, Deauville at response assessment

Consensus read:
discrepant sites classified as:
- Perceptual error
- Technical error

Analysis 2
Repeat analysis after correction for perceptual error

PET/MRI in pediatric Hodgkin lymphoma
FIGURE 2

Axial CT (A), $^{18}$F-FDG PET (B) and fused $^{18}$F-FDG PET/CT (D) images showing right upper deep cervical lymphadenopathy (arrows) in a 12 year old male with lymphocyte predominant Hodgkin lymphoma at staging. Axial T2 HASTE MRI (D), $^{18}$F-FDG PET (E) and fused $^{18}$F-FDG PET/MRI (F) showing the same.
FIGURE 3

A)B)C) Show coronal STIR, $^{18}$F-FDG PET and $^{18}$F-FDG PET/MR fused images demonstrating right supraclavicular fossa and right paratracheal lymphadenopathy, and a right lung nodule (arrows), in a 14 year old male with Hodgkin lymphoma.

D)E)F) Early response assessment following 2 cycles of chemotherapy shows residual, smaller right supraclavicular fossa lymphadenopathy (on D, coronal STIR), with no uptake (on E and F, coronal $^{18}$F-FDG PET and fused $^{18}$F-FDG PET/MR (arrows)). The other sites of disease have resolved.
PET/MRI in pediatric Hodgkin lymphoma