

Uniwersytet Jagielloński
Collegium Medicum

Anna Kocurek

**Sposoby wczesnej oceny skuteczności leczenia
zaawansowanej postaci ziarnicy złośliwej (HL) w
badaniu PET/CT.**

Praca doktorska

Promotor: prof. dr hab. n. med. Wojciech Jurczak

Pracę wykonano w Zakładzie Dydaktyki Medycznej UJ CM

Kierownik jednostki: dr n. med. Grzegorz Cebula

Kraków, rok 2020

Szczególne podziękowania kieruję do osób, bez pomocy których niniejsza praca by nie powstała:

prof. dr hab. n. med. Wojciecha Jurczaka, jej głównego pomysłodawcy i promotora,

dr hab. n. med. Bogdana Małeckiego, który udostępnił badania krakowskich pacjentów, wykonane w latach 2006-2011 w Zakładzie Medycyny Nuklearnej Centrum Onkologii w Bydgoszczy,

dr n. chem. Elżbiety Broniatowskiej, która wskazała kierunek analizy statystycznej,

oraz wszystkich innych, którzy nie przestawali we mnie wierzyć i mnie wspierać.

SPIS TREŚCI

Wstęp.....	7
Cel badania.....	9
Ograniczenie zakresu rozważań.....	9
Streszczenie badania.....	10
Uzyskane wyniki	11
Przegląd literatury	14
Ziarnica złośliwa – wprowadzenie	14
Ocena stopnia zaawansowania choroby (klasyfikacja Ann Arbor).....	15
Ocena rokowania – IPS.....	16
Ocena odpowiedzi na leczenie – kryteria IWG, zwane kryteriami Chesona.....	16
¹⁸ F-FDG PET oraz kryteria jego oceny – IHP, Deauville, Lugano i LYRIC	18
Kryteria RECIL	20
Terapia I rzutu ziarnicy złośliwej.....	21
Terapie celowane w leczeniu I rzutu.....	23
Badanie interim PET/CT i standard jego oceny	24
Ocena parametrów anatomicznych w badaniu PET/CT	25
Ocena ilościowych parametrów PET - SUV zmiany i jego pochodne	26
Metody oceny półautomatycznej - MTV i TLG	27
Materiały i metody	30
Metodyka i dobór badanej grupy	31
Wyjściowa ocena badania PET/CT	32
Parametry anatomiczne badania PET/CT	33
Parametry PET	35
Pomiary półautomatyczne	36
Analiza statystyczna	37
Wyniki	39
Opis badanej próby	39
Wyjściowa ocena badania PET/CT	44
Ocena parametrów anatomicznych badania PET/CT	45
Ocena parametrów związanych z SUV.....	50
Dyskusja	57

Interpretacja wczesnej odpowiedzi na leczenie - badania interim PET/CT	61
Włączenie do oceny parametrów z anatomicznej części badania PET/CT.....	62
Parametry badania PET – standardowe kryteria oceny	67
Ilościowe metody interpretacji PET	70
Masa guza (ang. tumour burden) oceniana przy pomocy metod półautomatycznych w badaniu PET – metaboliczna objętość guza (MTV) oraz całkowita glikoliza zmiany (TLG)	75
Ograniczenia niniejszego badania oraz potencjalne przeszkody we wdrożeniu proponowanych rozwiązań:.....	80
Podsumowanie	83
Perspektywa kontynuacji badań	86
Piśmiennictwo:	88
Aneks 1: Zbiorcze zestawienie klasyfikacji odpowiedzi na leczenie w chłoniakach.....	106
Spis tabel.....	109
Spis rycin.....	111
Streszczenie w języku polskim	112
Streszczenie w języku angielskim.....	113

WSTĘP

Ziarnica złośliwa (chłoniak Hodgkina, ang. HL – Hodgkin Lymphoma) nawet w zaawansowanym stadium jest obecnie nowotworem o dobrym rokowaniu. Stosowany już od ponad 40 lat schemat chemioterapii ABVD (doksorubicyna, etopozyd, winblastyna i dakarbazyna) (Bonadonna *et al.*, 1975), pozwala na wyleczenie 70 do 80% chorych (Johnson *et al.*, 2016). Większą skuteczność może mieć eskalowany BEACOPP (eskBEACOPP), schemat z bleomycyną, etopozydem, doksorubicyną, cyklofosfamidem, winkrystyną, prokarbazyną i prednizonem, w którym dawki etopozydu, doksorubicyny i cyklofosfamidu są wyższe od standardowych (Engert *et al.*, 2009; Mounier *et al.*, 2014). Wyższy o 5-10% odsetek chorych, żyjących bez progresji choroby w 5 lat po podaniu 5-8 cykli eskalowanego BEACOPP (Skoetz *et al.*, 2013), oznaczał dla części z nich trudne do zaakceptowania efekty działań niepożądanych, takie jak niepłodność czy wtórny do leczenia zespół mielodysplastyczny lub wręcz ostrą białaczkę szpikową (Behringer *et al.*, 2013; Kreissl *et al.*, 2016).

Próby utrzymania wysokiej skuteczności leczenia eskalowanym BEACOPP przy zmniejszeniu częstości skutków ubocznych są obecnie realizowane przy pomocy dwóch strategii – obu, opartych na wczesnej ocenie skuteczności leczenia. W części badań klinicznych leczenie jest rozpoczynane przy pomocy schematu ABVD, a eskalację chemioterapii do eskBEACOPP proponuje się jedynie chorym, którzy nie uzyskują regresji zmian po pierwszych 2 cyklach. W ośrodku Krakowskim, chorych z HL z dużą masą lub zaawansowanym stadium klinicznym, leczono od początku eskBEACOPP, co pozwalało na przełamanie pierwotnej oporności nowotworu na leczenie, zmniejszając intensywność leczenia po 2 cyklach – przechodząc do schematu leczenia ABVD (Casasnovas *et al.*, 2019).

Podstawą do podejmowania decyzji klinicznych w chłoniaku Hodgkina jest wczesna ocena odpowiedzi (ang. interim) po 2 cyklach chemioterapii przy pomocy pozytonowej tomografii emisyjnej (ang. PET – Positron Emission Tomography), wykonywanej z użyciem 2-^[18F]fluoro-2-deoxy-D-glukozy (FDG). Badanie to jest najczęściej wykonywane jednocześnie z tomografią komputerową (TK) jako PET/CT

(ang. Positron Emission Tomography/Computer Tomography¹). Jego wynik okazał się być najlepszym wskaźnikiem prognostycznym, od którego zależy zarówno mediana długości czasu przeżycia bez progresji (ang. PFS – Progression Free Survival) jak i przeżycia całkowitego (ang. OS – Overall Survival) (Adams, Nievelstein and Kwee, 2015; Kanoun, Rossi and Casasnovas, 2018). Wczesna ocena odpowiedzi w PET/CT charakteryzuje się umiarkowaną czułością i dość wysoką swoistością (odpowiednio 70,8% [95% przedział ufności (ang. CI – Confidence Interval): 64,7-76,4%] oraz 89,9 % [95% CI: 88-91,6%]) (Adams, Nievelstein and Kwee, 2015).

Opublikowane badania kliniczne cechują się dużą różnorodnością m. in. w sposobach interpretacji wczesnego PET/CT. Podstawowy sposób oceny tego badania jest ograniczony do części PET i bazuje na pięciostopniowej skali (ang. 5-PS, 5-Point Scale), znanej powszechnie jako Deauville, w której SUVmax (maksymalna standaryzowana wartość wychwytu, ang. maximum Standardized Uptake Value) zmiany jest porównywany w stosunku do otoczenia, puli krwi narządów śródpiersia oraz wątroby. PET/CT oceniane wg skali Deauville ma wysoką wartość predykcyjną, jeśli chodzi o identyfikację chorych nie odpowiadających na leczenie, mniejszą dla identyfikacji chorych o dobrej odpowiedzi klinicznej. Podkreśla się konieczność dalszych badań pozwalających na ocenę znaczenia prognostycznego wczesnej oceny PET/CT i sposobów jego interpretacji (Adams and Kwee, 2016).

W opublikowanej już pracy, autorka sprawdziła przydatność łącznej oceny regresji morfologicznej (dane z części tomograficznej) i metabolicznej (dane z części nuklearnej badania PET/CT) u pacjentów z chłoniakiem pierwotnym śródpiersia z komórek B (ang. PMBCL – Primary Mediastinal large B-Cell Lymphoma) (Kocurek *et al.*, 2016). Uzyskane wyniki były na tyle obiecujące, że podjęto się opracowania materiału obejmującego większą i bardziej różnorodną grupę chorych. Przedstawiona w niniejszej rozprawie ocena przydatności badania PET/CT, służącego jako wczesna ocena odpowiedzi na leczenie u chorych z ziarnicą złośliwą, stała się częścią większej analizy

¹ Przyjętym w języku polskim skrótem, oznaczającym tomografię komputerową, jest TK. Jednakże w używanym powszechnie skrócie PET zachowano składnię angielską – dlatego też w niniejszej pracy zdecydowano się na używanie całości określenia pochodzącego z języka angielskiego - PET/CT.

dotyczącej indywidualizacji leczenia u chorych w zaawansowanym stadium klinicznym, leczonych w Klinice Hematologii UJ CM.

Cel badania

Celem pracy była identyfikacja parametrów wczesnej oceny FDG-PET/CT, umożliwiających w najlepszy sposób wskazanie chorych z dobrą odpowiedzią na pierwsze 2 cykle eskBEACOPP, u których zmniejszenie intensywności dalszej chemioterapii, nie obniży skuteczności leczenia.

Ograniczenie zakresu rozważań

Ziarnica złośliwa jest stosunkowo częstym podtypem chłoniaka, na temat którego powstało wiele publikacji (ogólne zapytanie „Hodkin lymphoma” w bazie PubMed w dniu 20.06.2020 zwróciło 74 045 wyników). Z tego powodu, konieczne było ściśle zawężenie poszukiwań literaturowych oraz obszaru badań. Kryteria włączenia do badania ograniczono do dorosłych chorych z ziarnicą złośliwą w zaawansowanym stadium klinicznym, wyjściowo leczonych w Klinice Hematologii UJCM przy pomocy schematu eskBEACOPP (z możliwą następną deeskalacją leczenia) i kontrolowanych w badaniu PET/CT przynajmniej w 2 punktach czasowych (przed rozpoczęciem leczenia i jako wczesna ocena skuteczności po 2 cyklu chemioterapii). Najważniejszymi punktami końcowymi niniejszego badania było niepowodzenie leczenia I rzutu (rozumiane jako pierwotna oporność lub wznowa procesu), oraz czasy przeżycia pacjentów (PFS i OS).

Podobnie zawężono poszukiwania literaturowe do zagadnień ściśle związanych z tematyką pracy – skupiono się na ogólnej charakterystyce choroby u osób dorosłych i ocenie odpowiedzi na leczenie I rzutu w badaniach obrazowych, natomiast nie uwzględniono w nich przykładowo szczegółowych zagadnień histologicznych, problematyki związanej z leczeniem choroby odpornej lub nawrotowej czy odrębności choroby w populacji pediatrycznej.

Streszczenie badania

Badanie retrospektywne obejmujące zebranie danych na podstawie klinicznej dokumentacji medycznej z lat 2006 – 2011 (historii chorób) oraz szczegółowych pomiarów parametrów PET i tomograficznych w obrębie dostępnych badań PET/CT z okresu leczenia. Uzyskane dane pozwalają na ocenę w długim okresie obserwacji (mediana obserwacji 10,2 lat).

Kryteria włączenia: pacjenci z zaawansowaną postacią ziarnicy złośliwej (Ann Arbor II BX, III i IV), leczeni w latach 2006 – 2011 w Klinice Hematologii SU w Krakowie przy pomocy schematu chemioterapii eskBEACOPP, w razie dobrej odpowiedzi klinicznej zmienianego na ABVD, oceniani w badaniu PET/CT przynajmniej w 2 punktach czasowych (przed rozpoczęciem leczenia oraz jako wczesna ocena skuteczności po 2 cyklu chemioterapii).

Kryteria wykluczenia: brak dostępności dokumentacji medycznej i/lub badań PET/CT

Badana populacja została podzielona na dwie grupy: pacjenci z pierwotną opornością lub wznową choroby po leczeniu I rzutu oraz ci, którzy odpowiedzieli na leczenie I rzutu i żyli bez wznowy choroby przynajmniej 5 lat od rozpoznania. W skład badania weszła, poza uwzględnieniem danych klinicznych, powtórna ocena badań PET/CT wykonywanych przed leczeniem oraz wcześniej podczas leczenia. Uwzględniała ona ocenę wartości rokowniczej informacji anatomicznych z części tomograficznej badania (z wykorzystaniem różnych metod pomiaru, w tym wolumetrycznych) oraz jakościowej i ilościowej oceny części PET (w tym półautomatycznych metod oceny). W analizie części tomograficznej badania wykorzystywano następujące parametry: SLD (ang. Sum of Longest Diameters – suma najdłuższych wymiarów zmian), SPD (ang. Sum of the Products of the longest Perpendicular Diameters – suma iloczynów prostopadłych wymiarów zmian chłoniakowych) oraz objętość zmian przed i po rozpoczęciu leczenia (ang. V – Volume). W analizie PET wykorzystano parametry pośrednio oceniające aktywność metaboliczną guza (SUVmax i SUVpeak), narządów referencyjnych (pęka krwi narządów śródpiersia i wątroba) oraz półautomatyczne metody pozwalające na

określenie metabolicznej objętości guza (ang. MTV – Metabolic Tumour Volume) oraz całkowitej glikolizy w obrębie zmiany (ang. TLG – Total Lesion Glycolysis).

Uzyskane wyniki

Kryteria włączenia do badania spełniło 58 pacjentów. W momencie rozpoczęcia leczenia byli w wieku od 23 do 36 lat (mediana 29 lat), mężczyźni stanowili 38,3% populacji badanej (23 osoby). Wszyscy zostali ocenieni w badaniu PET/CT wykonanym w jednym ośrodku wg. tego samego schematu, zazwyczaj po drugim cyklu terapii (u 9 osób po trzecim). Dalsze leczenie było dostosowywane do uzyskanej odpowiedzi: u chorych z całkowitą remisją (ang. CR – Complete Remission), w badaniu PET/CT (Deauville 1-3) leczenie było deeskalowane do ABVD (n = 45). U chorych bez CR (n = 15), kontynuowano eskBEACOPP, w 4 przypadkach poddając pacjentów konsolidacji wysokodawkowaną chemioterapią wspomaganą przeszczepieniem komórek macierzystych (ang. ASCT – Autologous Stem Cell Transplant). Uzupełnieniem chemioterapii I rzutu, była u 21 chorych radioterapia miejsc pierwotnie zmienionych (ang. IFRT – Involved Field RadioTherapy).

W całej grupie u 50 osób osiągnięto wieloletnią remisję (mediana odpowiedzi 10,2 lat). U 8 pacjentów doszło do progresji lub wznowa choroby: trzy z nich zmarły w wyniku progresji, dalsze dwie zmarły w trakcie leczenia z innych powodów (ze względu na brak możliwości rzetelnej oceny wyników leczenia wyłączono ich z analizy).

W wykonanej analizie danych z części anatomicznej badania PET/CT udowodniono, że informacje z wczesnego badania PET/CT (zarówno SPD, jak i suma objętości zmian) pozwalają podobnie jak kryteria Deauville ocenić rokowanie chorych (wartości p jednoczynnikowej regresji logistycznej wynoszą 0,01 dla kryteriów Deauville, 0,02 dla SPD i 0,002 dla sumy objętości). Połączenie kryteriów PET i tomograficznych pozwoliło zwiększyć dodatnią wartość predykcyjną (ang. PPV – Positive Predictive Value) przy zachowaniu wysokiej ujemnej wartości predykcyjnej badania (ang. NPV – Negative Predictive Value) – wartości wyniosły odpowiednio 86% i 96,1% v porównaniu do 50% i 97,7% przy kryteriach PET.

W szczegółowej analizie wyników części nuklearnej badania PET/CT porównano różne metody oceny ilościowej tego badania. Zaproponowano wykorzystanie w ocenie chorych z ziarnicą złośliwą parametru ilorazu SUVmax zmiany i wątroby (rPET), który do tej pory nie był analizowany w tej postaci choroby. Jest to parametr bardzo prosty do wyliczenia – bazuje na zmiennych standardowo wyznaczanych przy ocenie badania PET/CT. Przyjęcie punktu odcięcia na poziomie 1,3 dla tego parametru wyliczonego dla badania PET wykonywanego po pierwszych cyklach chemioterapii (interim PET) pozwala na osiągnięcie wysokiej swoistości (96% vs. 86% przy standardowych kryteriach PET) i dokładniejsze określenie grupy pacjentów z korzystnym wynikiem leczenia I rzutu. Parametrem podobnie łatwym do wykorzystania oraz pozwalającym na dobre zróżnicowanie chorych, u których wystąpiło niepowodzenie leczenia I rzutu, jest ΔSUVmax , opisujący różnicę w poziomie wychwytu FDG pomiędzy badaniem przed i w trakcie leczenia i dostarczającym istotnych wskazówek klinicznych na temat odpowiedzi metabolicznej guza na leczenie. W ramach sprawdzenia przydatności półautomatycznej oceny badania PET/CT wykazano szczególną wartość MTV i TLG wyliczonych dla badania interim przy pomocy punktu odcięcia odnoszącego się do SUV wątroby ($\text{SUV}_{\text{mean}+3\text{SD}}$).

Uzyskane wyniki są ostateczną analizą po latach obserwacji, aktualnie w fazie ostatecznego przygotowania publikacji w recenzowanych czasopismach medycznych.

Tabela 1: Zestawienie aktualnego stanu wiedzy oraz innowacji wnoszonych przez niniejszą pracę.

Co wiadomo na temat poruszanego zagadnienia:	Co nowego wnosi praca:
<ol style="list-style-type: none"> 1. Leczenie przy pomocy schematu eskBEACOPP charakteryzuje się wysoką skutecznością, ale jest obarczone trudnymi do zaakceptowania skutkami ubocznymi. 2. Rozpoczęcie leczenia od eskBEACOPP z możliwością deeskalacji do ABVD u pacjentów z CR we wczesnej ocenie skuteczności leczenia pozwala na porównywalną skuteczność leczenia przy zminimalizowaniu jego skutków ubocznych. 3. Wczesna ocena w badaniu PET/CT z wykorzystaniem standardowych metod ogranicza się do oceny części PET badania i charakteryzuje się wysoką czułością oraz niską swoistością. 	<ol style="list-style-type: none"> 1. Poprawa swoistości wczesnej oceny PET/CT przez dołączenie do informacji o metabolizmie guza danych anatomicznych. 2. Optymalizacja wykorzystywanych metod oceny metabolizmu nowotworu na wczesnym etapie leczenia. 3. Propozycje optymalizacji protokołów klinicznych pozwalających podjąć decyzję o deeskalacji leczenia.

Ziarnica złośliwa – wprowadzenie

Ziarnica złośliwa stanowi mniej niż 1% wszystkich rozpoznań nowotworów złośliwych (w tym około 10% chłoniaków), z rocznym występowaniem z częstością 2-3 przypadki na 100 000 w świecie zachodnim (Townsend and Linch, 2012; Siegel *et al.*, 2014). Jej leczenie, wyjściowo bazujące na radioterapii, następnie na kombinacji różnych chemioterapeutyków z dodatkiem lub bez radioterapii, obecnie także z dodatkiem przeciwciał monoklonalnych, całkowicie zmieniło obraz tej choroby na przestrzeni ostatnich lat (Johnson, 2016). Obecnie pacjenci we wczesnej postaci choroby mogą liczyć na wyleczenie w 95% przypadków, natomiast w przypadku zaawansowanej choroby potencjalny wskaźnik wyleczenia może przekraczać 90% (Johnson, 2016). Rosnące prawdopodobieństwo wyleczenia w grupie pacjentów, którzy najczęściej są młodymi dorosłymi, zmieniło punkt ciężkości terapii – z pierwotnego skupienia się na eradykacji choroby na bardziej kompleksowe podejście, łączące szanse na wyzdrowienie z obniżaniem ryzyka skutków ubocznych (szczególnie długoterminowych). Wyniki badań na dużych populacjach wskazują, że obecnie zarówno we wczesnej (Radford *et al.*, 2015), jak i zaawansowanej postaci choroby (Johnson *et al.*, 2016) ryzyko zgonu z powodu choroby obecnie jest mniejsze, niż ryzyko zgonu z innych powodów, spośród których są toksyczne skutki terapii.

Mediana wieku chorych na ziarnicę złośliwą wynosi 28 lat, drugi szczyt zachorowań jest notowany pomiędzy 60 a 70 rokiem życia (Matasar and Zelenetz, 2008; Townsend and Linch, 2012). Typowo wyjściowym objawem choroby jest niebolesna limfadenopatia, często w lokalizacji szyjnej lub nadobojczykowej. Ponad połowa pacjentów ma masę w śródpiersiu, która może być bezobjawowa lub powodować duszność, kaszel lub zwężenie żyły głównej górnej (Mauch *et al.*, 1993). Objawy ogólne występują u 25% pacjentów. Gorączka, zlewne poty nocne lub utrata ponad 10% masy ciała w przeciągu 6 miesięcy są określane jako objawy B i mają znaczenie prognostyczne.

Obecnie nie uważa się, aby inne objawy, takie jak świąd skóry, męczliwość i ból związany ze spożyciem alkoholu miały znaczenie prognostyczne (Townsend and Linch, 2012).

Ocena stopnia zaawansowania choroby (klasyfikacja Ann Arbor)

Wyjściowe określenie poziomu zaawansowania choroby (ang. staging) jest dokonywane przy pomocy systemu Ann Arbor. Wyjściowo klasyfikacja ta, wprowadzona w roku 1971, odnosiła się do pól włączanych do radioterapii (Carbone *et al.*, 1971). System został zmodyfikowany w 1989 r. poprzez wprowadzenie modyfikacji Costwolds, uwzględniającej dane z badania tomografii komputerowej (Lister *et al.*, 1989). Ten sposób oceny bazuje na rozległości zajętych grup węzłowych oraz narządów pozawęzłowych i jest wciąż wykorzystywany dla oceny rokowania oraz planowania leczenia (szczegóły zostały przedstawione w tabeli 2):

Tabela 2: Stopień wyjściowego zaawansowania choroby Ann Arbor (za Szczeklik *et al.*, 2018, p. 1673, zmieniono).

Ann Arbor	Charakterystyka
I	zajęcie pojedynczej grupy węzłowej lub jednego narządu pozalimfatycznego (I E)
II	Zajęcie 2 lub więcej grup węzłowych po tej samej stronie przepony lub umiejscowione (jednoogniskowe) zajęcie narządu pozalimfatycznego i 2 lub więcej grup węzłów chłonnych po tej samej stronie przepony (II E)
III	zajęcie węzłów chłonnych po obu stronach przepony, czemu towarzyszyć może jednoogniskowe zajęcie narządu pozalimfatycznego (III E) lub zajęcie śledziony (III S), lub zajęcie narządu pozalimfatycznego (jednoogniskowe) i śledziony (III SE)
IV	Rozsiane zajęcie narządu pozalimfatycznego, niezależnie od stanu węzłów chłonnych
Dodatkowe oznaczenie wprowadzone na konferencji w Costwolds: X – duże zmiany węzłowe (ang. bulky disease), np. guz śródpiersia zajmujący > 1/3 szerokości kłp. lub masa węzłowa w innej okolicy o największym wymiarze ≥ 10 cm	

Ocena rokowania – IPS

W roku 1998, na podstawie wielośrodkowych wyników leczenia 5141 chorych z zaawansowaną postacią ziarnicy złośliwej, Hasenclever i Diehl zaproponowali wprowadzenie ogólnie dziś przyjętego wskaźnika rokowania – nazwanego następnie Międzynarodowym Wskaźnikiem Prognostycznym (International Prognostic Score) lub wskaźnikiem Hasenclevera (Hasenclever and Diehl, 1998). Jest on liczony przy uwzględnieniu siedmiu czynników klinicznych i laboratoryjnych (tabela 3). Przy wykorzystaniu tradycyjnych terapii obecność każdego z nich obniżała ogólne 5-letnie przeżycie wolne od choroby o od 7 do 10% (Townsend and Linch, 2012). Wskaźnik ten został także przyjęty do oceny oraz ewentualnej deintensyfikacji leczenia u chorych z wczesnymi postaciami HL (Engert *et al.*, 2010).

Tabela 3: Niekorzystne czynniki rokownicze w ziarnicy złośliwej – IPS.

Wiek > 45 lat
Płeć męska
Stężenie albuminy w surowicy < 40g/L
Stężenie hemoglobiny < 105 g/L
Stopień IV choroby
Leukocytoza ($\geq 15 \times 10^9$ leukocytów/L)
Limfopenia (< $0,6 \times 10^9$ limfocytów/L, lub < 8% całkowitej liczby leukocytów)

Ocena odpowiedzi na leczenie – kryteria IWG, zwane kryteriami Chesona

Wyjściowo kryteria oceny w chłoniakach były wykorzystywane jedynie do oceny odpowiedzi na leczenie w chwili zakończenia leczenia. Różniły się one znacząco

pomiędzy ośrodkami klinicznymi i grupami badającymi te choroby i brały pod uwagę liczne parametry. Ich przykładami mogą być: obecność objawów klinicznych, wyniki badań laboratoryjnych, wyniki zwiadowczej laparotomii/reoperacji i badań obrazowych, takich jak zdjęcia rtg, scyntygrafia z wykorzystaniem galu czy tomografia komputerowa (Adams and Kwee, 2017).

Pierwszą próbą standaryzacji na dużą skalę kryteriów oceny były kryteria Międzynarodowej Grupy Roboczej (ang. IWG, International Working Group), opublikowane w roku 1999 i będące wynikiem współpracy grupy klinicystów, radiologów i anatomów (Cheson *et al.*, 1999). Zalecenia te zostały szeroko przyjęte w świecie przez klinicystów oraz agencje regulacyjne. Wyjściowo odnosiły się do chłoniaków nieziarniczych (ang. n-HL – non-Hodgkin lymphoma), następnie zostały zaadoptowane także przez grupy badawcze koncentrujące się na leczeniu chorych z ziarnicą złośliwą (Cheson *et al.*, 2007). W tych kryteriach stan remisji był określany przez ocenę regresji węzłowej i pozawęzłowej choroby, wykonywaną na bazie badania fizykalnego, biopsji szpiku kostnego oraz tomografii komputerowej. Bazując na zmianie wielkości zmian nowotworowych wyróżniono pięć kategorii odpowiedzi na leczenie:

- progresja choroby (ang. PD – Progressive Disease),
- stabilizacja choroby (ang. SD - Stable Disease),
- częściowa remisja (ang. PR – Partial Remission),
- całkowita remisja (ang. CR – Complete Remission) oraz
- niepotwierdzona całkowita remisja (ang. CRu – Complete Remission unconfirmed).

Szczególnie ta ostatnia kategoria była źródłem licznych nieporozumień i rozbieżności w ocenach różnych lekarzy i badaczy. Była ona przypisywana pacjentom ze znacznym zmniejszeniem wielkości guza w badaniu wykonywanym na zakończenie leczenia, lecz z obecnością masy rezidualnych w badaniu tomografii komputerowej. Jednak obecność takiej masy w chłoniakach – szczególnie ziarnicy złośliwej – po zakończeniu leczenia może odpowiadać zarówno tkance nowotworowej, jak i włóknistej. Metodą, którą zaproponowano oraz szeroko wprowadzono dla rozróżnienia tych dwóch sytuacji, stała

się pozytronowa tomografia emisyjna (PET), wykonywana przy pomocy ^{18}F -fluoro-2-deoxy-D glukozy (^{18}F -FDG).

^{18}F -FDG PET oraz kryteria jego oceny – IHP, Deauville, Lugano i LYRIC

Badanie ^{18}F -FDG PET, jeszcze na przełomie milenium uważane przez wielu za ciekawostkę o zastosowaniu głównie w badaniach naukowych, stało się potężnym narzędziem wykorzystywanym zarówno przy diagnostyce nowotworów (szczególnie cechujących się wysokim poziomem metabolizmu glukozy – z ang. FDG-avid), ocenie wyjściowej stopnia zaawansowania, jak i w różnych momentach oceny odpowiedzi na leczenie ziarnicy złośliwej. Dla ułatwienia interpretacji obrazów i zwiększenia czułości oraz swoistości badania, jest ono najczęściej łączone z tomografią komputerową, wykonywaną bez lub z dożylnym podaniem kontrastu.

Okazało się, że połączenie tych badań prowadzi do zwiększenia czułości i swoistości oceny zajęcia węzłów chłonnych do odpowiednio 94% i 100%, podczas gdy sama tomografia komputerowa, wykonywana po dożylnym podaniu kontrastu, cechuje się 88% czułością i 86% swoistością (Schaefer *et al.*, 2004). W tym samym badaniu dla zmian pozawęzłowych FDG PET/CT wykazało czułość na poziomie 88% i swoistość 100%, w sytuacji, gdy dla tomografii te same wartości wyniosły odpowiednio 50% i 90%. Przy wykorzystaniu tego badania do oceny wyjściowej (ang. staging) w ziarnicy złośliwej, stwierdzany stopień zaawansowania był podwyższany u 32% chorych i obniżany u 15%, co prowadziło do zmiany postępowania klinicznego u 33% pacjentów (Paes *et al.*, 2010). Znaczący wpływ na zmianę stageru pacjentów wykazali też Bednaruk-Młyński *et al.* (2015).

Wyjściowo badania PET/CT również były interpretowane w sposób niewystandaryzowany, co zmieniło się w roku 2007, po ustaleniu i przyjęciu kryteriów Międzynarodowego Projektu Harmonizacyjnego (ang. IHP – International Harmonization Project) (Cheson *et al.*, 2007; Juweid *et al.*, 2007), zwanych też zrewidowanymi kryteriami Chesona. Zgodnie z tymi kryteriami wychwyty intensywniejszy

od aktywności puli krwi narządów śródpiersia w zmianach większych niż 2 cm oraz wyższy niż otoczenie w zmianach mniejszych wskazuje na obecność przetrwałej choroby.

Kolejnym ważnym krokiem w interpretacji badania PET/CT stało się włączenie do badań klinicznych wizualnej oceny wychwytu zmian nowotworowych w skali pięciostopniowej, zwanej też londyńską (miało to miejsce w Deauville, we Francji, stąd wzięła się też nazwa tych kryteriów) (Meignan *et al.*, 2009). W skali tej punktami odcięcia dla kolejnych stopni stało się porównanie wychwytu zmian nowotworowych do: otoczenia, puli krwi narządów śródpiersia (ang. MBP – Mediastinal Blood Pool lub MBPS – Mediastinal Blood Pool Structures) oraz wątroby, od której może być nieznacznie lub znacząco wyższy (pojawienie się nowych zmian również oznacza najwyższy punkt w skali). Kryteria te następnie włączono do klasyfikacji Lugano (Barrington *et al.*, 2014; Cheson *et al.*, 2014), w której generalnie dla chłoniaków wykazujących się wysokim metabolizmem glukozy, zaproponowano poziom wychwytu radioznacznika przez wątrobę jako punkt odcięcia dla zdrowych i chorych, szczególnie w badaniu interim, wykonywanym w trakcie chemioterapii (i.e. punkty 1-3 w skali pięciostopniowej są uważane za wyznaczniki całkowitej odpowiedzi na leczenie, z kolei punkty 4-5 – niepowodzenia leczenia) (Adams and Kwee, 2017), z zastrzeżeniem, że to podejście może zostać zastrzeżone w przypadku konkretnych wskazań klinicznych.

Szczegółowe porównanie ww. skal zawarto w Aneksie 1.

Ze względu na rosnącą rolę immunomodulujących biofarmaceutyków w leczeniu chłoniaków, konieczne okazało się rozszerzenie klasyfikacji Lugano. Leki te mogą mieć znacznie bardziej zróżnicowany wpływ na zachowanie nowotworu niż klasyczne chemioterapytyki. W szczególności opisano występowanie wyjściowego zwiększenia się wielkości i ilości zmian, wyprzedzającego aktualną odpowiedź na leczenie (zjawisko to nosi nazwę pseudoprogresji). Ponadto zaobserwowano „zjawisko flary” (flare phenomenon), tj. zwiększony wychwyt radioznacznika, występujący jako część (pozytywnej) odpowiedzi na leczenie (Skusa *et al.*, 2020). Dlatego też zaproponowano dołączenie do kategorii PD (zgodnej z Lugano) podgrupy „odpowiedź niezdeterminowana” (ang. IR – Indeterminate Response). Umożliwia ona początkową kontynuację leczenia, po której powinna być wykonana ponowna ocena w badaniach obrazowych w terminie do 12 tygodni, potrzebna do odpowiedzi na pytanie, czy w

danym przypadku wystąpiła rzeczywista progresja, czy też zmiany związane z leczeniem. Kryteria te otrzymały nazwę modyfikacji LYRIC (ang. Lymphoma Response to Immunomodulatory Therapy Criteria) (Cheson *et al.*, 2016).

Kryteria RECIL

W przedstawionych powyżej kryteriach oceny odpowiedzi na leczenie w ziarnicy złośliwej brano pod uwagę albo wyłącznie poziom wychwytu ¹⁸F-fluoro-2-deoxy-D glukozy przez zmiany nowotworowe albo (gdy np. badanie PET było niedostępne) zmianę ich wielkości, liczoną na podstawie zmiany SPD – sumy iloczynów wymiarów prostopadłych zmian nowotworowych (wyznacznik ten wprowadzono w kryteriach IWG oraz zachowano w klasyfikacji Lugano).

Równocześnie – jak powyżej wspomniano – zauważono, że szereg czynników, w tym leki nowej generacji, moderują metabolizm guzów, wychwyty glukozy, stan zapalny w otoczeniu nowotworu, a nawet mogą powodować pseudoprogresję guza – i tym samym mogą potencjalnie zwiększać ilość fałszywie pozytywnych lub fałszywie negatywnych wyników badania FDG-PET (Meignan *et al.*, 2015).

Dlatego też w zaproponowanej stosunkowo niedawno klasyfikacji RECIL (Response Evaluation Classification for Lymphoma - Younes *et al.*, 2017) zaproponowano połączenie zarówno oceny badania PET/CT (bazującej – jak w klasyfikacji Lugano - na kryteriach Deauville) oraz oceny morfologii zmian. Wzorując się na kryteriach oceny obowiązujących dla nowotworów litych, RECIST (Eisenhauer *et al.*, 2009), zbadano możliwość wykorzystania pomiarów pojedynczego wymiaru zmian węzłowych. Finalnie, przy występowaniu u chorego mnogich zmian, zaproponowano wykorzystanie nowego parametru – sumy najdłuższych wymiarów do trzech zmian węzłowych (SLD).

Nowe kryteria oceny odpowiedzi na leczenie w chłoniakach zostały zaaprobowane jak do tej pory głównie do oceny chłoniaków nieziarnicznych (Younes *et al.*, 2017).

Terapia I rzutu ziarnicy złośliwej

Zaawansowana ziarnica złośliwa jest leczona przy pomocy schematów uwzględniających równoczesne podanie różnych chemioterapeutyków. Podstawowym schematem leczenia I rzutu stała się wprowadzona przeszło 40 lat temu kombinacja ABVD (doksorubicyna, etopozyd, winblastyna i dakarbazyna – Bonadonna *et al.*, 1975). Pozwala ona na osiągnięcie od 70% do 80% trwałych remisji choroby - odsetka zbliżonego do wcześniej stosowanych bardziej złożonych schematów leczenia i wiążących się z większą toksycznością (Canellos *et al.*, 1992; Viviani *et al.*, 1996; Duggan *et al.*, 2003; Johnson *et al.*, 2005).

Jedną z modyfikacji tego schematu, bazująca na połączeniu różnych sposobów leczenia, została wprowadzona przez grupę związaną z Uniwersytetem Stanford. W schemacie Stanford V próbowano zredukować kardio- i pulmonotoksyczność leczenia, wykorzystując zredukowane dawki doxorubicyny i bleomycyny. Skrócono też okres leczenia do 12 tygodni. Jednak dla zachowania podobnych wyników leczenia, jak przy ABVD, konieczne okazało się włączenie napromieniania lokalizacji, w których wyjściowo obecne były ogniska choroby większe niż 5 cm oraz wyjściowo zajętej śledziony (Hoskin *et al.*, 2009; Gordon *et al.*, 2013) – co wiąże się z potencjalnymi skutkami ubocznymi radioterapii.

Kolejną próbą poprawy skuteczności leczenia ziarnicy stał się wprowadzony przez Niemiecką Grupę Badawczą Chłoniaka Hodgkina (ang. GHLSG – German Hodgkin Lymphoma Study Group) schemat eskalowany BEACOPP z bleomycynią, etopozydem, doksorubicyną, cyklofosfamidem, winkrystyną, prokarbazyną i prednizonem (w tym wyższymi od standardowych dawkami etopozydu, doksorubicyny i cyklofosfamidu - Engert *et al.*, 2009; Mounier *et al.*, 2014). Liczne badania kliniczne porównujące zastosowanie eskBEACOPP w stosunku do ABVD wykazały większy odsetek osób odpowiadających na terapię oraz dłuższy okres przeżycia wolnego od choroby, w tym odsetek przeżyć 5-letnich bez progresji choroby wyższych o 5-10 punktów procentowych (Skoetz *et al.*, 2013, 2017). Jednakże zastosowanie 6-8 cykli takiej chemioterapii dla części chorych wiąże się z występowaniem trudnych do

zaakceptowania efektów ubocznych. Pośród nich można wymienić trwałą bezpłodność, mielodysplazję oraz wtórne nowotwory - które mogą znacząco obniżyć różnicę w odległym przeżyciu (Behringer *et al.*, 2013; Kreissl *et al.*, 2016; Merli *et al.*, 2016). Wykazano, że 7 na 8 pacjentów poddanych takiej terapii niepotrzebnie doświadczyło jej powikłań – zostaliby wyleczeni przy pomocy standardowego schematu ABVD (Viviani *et al.*, 2011).

Leczenie zaawansowanych postaci ziarnicy złośliwej ewoluowało, od schematów chemioterapeutycznych opartych na wyższych dawkach i intensywności leczenia, do podejścia opartego na ocenie ryzyka. Milowym krokiem umożliwiającym zmianę tego podejścia były badania Gallaminiego i Hutchinsa (Hutchings *et al.*, 2005, 2006; Gallamini *et al.*, 2006), którzy – wyjściowo na niewielkich grupach pacjentów – pokazali, że badanie PET/CT wykonane wcześniej podczas leczenia (w badaniach wyjściowych po 2 lub 3 cyklu chemioterapii ABVD u chorych z zaawansowaną ziarnicą złośliwą) ma bardzo wysoką wartość prognostyczną, jeśli chodzi o przewidywanie wyniku leczenia. W dalszych badaniach zaobserwowano, że wartość prognostyczna tego badania przekracza dotychczas przyjęty IPS (Gallamini *et al.*, 2007). Z tego też powodu stało się ono czynnikiem, na którym oparto ewentualne modyfikacje leczenia w szeregu badań klinicznych, włączając w to zarówno stadia I-II HL (Raemaekers *et al.*, 2014; André *et al.*, 2017), jak i zaawansowane postaci tej choroby (Casasnovas *et al.*, 2015, 2019; Coyle, Kostakoglu and Evens, 2016; Borchmann, Goergen, *et al.*, 2017). W rzeczy samej, wczesna ocena odpowiedzi na leczenie w badaniu PET/CT otworzyła przed klinicystami nowe ścieżki postępowania klinicznego i stworzyła szansę na personalizację leczenia oraz zwiększyła możliwości precyzyjnego prowadzenia pacjentów (Subramaniam, 2017). Głównymi zaletami dla pacjentów są: poprawa wyników leczenia oraz zminimalizowanie skutków ubocznych terapii (Vassilakopoulos *et al.*, 2015). Ważnymi kierunkami badań klinicznych stała się eskalacja leczenia, najczęściej zaczynającego się od tradycyjnego schematu ABVD (przykładem może być badanie Johnson *et al.*, 2016), lub jego deeskalacja – najważniejszymi badaniami klinicznymi obierającymi ten kierunek stały się te, w których leczenie pacjentów rozpoczynano od schematu eskBEACOPP (Borchmann, Goergen, *et al.*, 2017; Casasnovas *et al.*, 2019).

Terapie celowane w leczeniu I rzutu

Połączenie przeciwciał z lekami cytotoksycznymi daje możliwość selektywnego dostarczenia do komórek nowotworu dawek leku, których nie dałoby się osiągnąć przy podaniu ogólnoustrojowym. Charakterystyczną cechą klasycznej ziarnicy złośliwej jest ekspresja receptora CD30 na komórkach Reed-Sternberga. Przykładem leku wykorzystującego ten mechanizm jest breximumab vedotin (BV), wyjściowo zarejestrowany dla opornej na leczenie postaci ziarnicy złośliwej (Shanbhag and Ambinder, 2018).

W badaniu I fazy u pacjentów z zaawansowaną ziarnicą złośliwą porównującym zastosowanie BV w połączeniu ze standardowym (ABVD) lub zmodyfikowanym (AVD) schematem leczenia, wykazano nieakceptowalnie wysoki odsetek powikłań pulmonologicznych (44%) w ramieniu ze schematem ABVD. Odsetek ten okazał się znacznie wyższy, niż w dotychczasowych badaniach z wykorzystaniem schematu ABVD - uznano, że za taki efekt było odpowiedzialne skojarzone podanie BV i bleomycyny, oraz że te dwa leki nie powinny być wspólnie podawane (Younes *et al.*, 2013).

Porównania schematu ABVD z AVD z dodatkiem BV (A+AVD) dokonano w wieloośrodkowym badaniu ECHELON-1, obejmującym 1334 pacjentów z Europy i USA w wieku powyżej 18 lat z III i IV stopniem zaawansowania ziarnicy złośliwej, wyjściowo nieleczonych z tego powodu (Connors *et al.*, 2018). W badaniu tym pacjenci zostali przebadani w badaniu interim PET po dwóch cyklach leczenia i ci, którzy uzyskali punktację 5 w skali 5-PS zostali uznani za osoby z progresją choroby i przekierowani do schematów leczenia II rzutu. Poza tym nie modyfikowano leczenia na podstawie interim PET – w standardzie pacjenci byli poddawani 6 cyklom chemioterapii. Po dwóch latach obserwacji (mediana 24,6 miesięcy) wyniki leczenia w grupie A-AVD okazały się nieco lepsze od tych w grupie ABVD – wskaźnik zmodyfikowanego 5-letniego przeżycia wolnego od choroby wyniósł 82% w stosunku do 77%. Zbliżony trend utrzymał się po trzech latach obserwacji (Straus *et al.*, 2020). W badanym schemacie wykazano większą częstość skutków ubocznych leczenia w stopniu 3 i 4, włączając w to gorączkę neutropeniczną i obwodową neuropatię. Z powodu powikłań hematologicznych

zmieniono protokół przez włączenie do leczenia w grupie AAVD czynnika wzrostu kolonii granulocytarnych (G-CSF) (Straus *et al.*, 2018). Zwracają uwagę bardzo złe wyniki leczenia wśród pacjentów z dodatnim wynikiem badania interim PET – w tej grupie 3 letni PFS wyniósł 58,2% lub jedynie 36,6% (odpowiednio ramię AAVD lub ABVD) (Lang and Crump, 2020).

Badanie interim PET/CT i standard jego oceny

W dotychczasowych badaniach kryterium dodatniego lub ujemnego interim PET/CT była przede wszystkim ocena w skali 5-PS, realizowana głównie poprzez porównanie wychwytu radioznacznika przez zmianę nowotworową do puli krwi narządów śródpiersia lub wątroby. Choć jest to uznany złoty standard, ujęty także w przyjętych klasyfikacjach odpowiedzi na leczenie w chłoniakach (Coyle, Kostakoglu and Evens, 2016), to jednak bazuje on głównie na rekomendacjach zespołów autorytetów. Przy takiej interpretacji – obowiązującej m. in. w kryteriach IHP i Lugano – rola badania tomograficznego w PET/CT u pacjentów z ziarnicą złośliwą sprowadza się korekcji atenuacji oraz ewentualnych podpowiedzi anatomicznych, np. przy wyznaczaniu SUV MBP (Barrington *et al.*, 2014).

Tymczasem przeszkodą w interpretacji tego badania może być fakt, że około 8-10% pacjentów z ziarnicą złośliwą prezentuje w badaniu interim PET niespecyficznie podwyższony wychwyt radioznacznika, najpewniej związany z tkankową odpowiedzią zapalną w reakcji na chemioterapię. W dużej mierze to oni tworzą grupę osób z fałszywie pozytywnymi wynikami tego badania (Gallamini and Zwarthoed, 2018). Potwierdza to wcześniejsze obserwacje, w których stwierdzono, że choć ujemna wartość predykcyjna (ang. NPV – negative predictive value) badania PET jest wysoka, jego dodatnia wartość predykcyjna (ang. PPV – positive predictive value) pozostaje umiarkowana (Elstrom, 2012).

Ponadto przy tak zdefiniowanej ocenie odpowiedzi na leczenie pomija się wiele płynących z badania PET/CT informacji, w tym informacji z części tomograficznej badania (np. poboczne znaczenie ma pomiar wielkości zmian oraz analiza ich wyglądu w TK) oraz

szeregu dodatkowych parametrów, opisujących część nuklearną tego badania. W dalszej części przeglądu literatury oraz przy prezentacji wyników niniejszej pracy przyjęto podział tych parametrów na trzy grupy:

1. wywodzące się z części tomograficznej badania PET/CT,
2. analiza wychwytu radioznacznika w zmianie/ach poprzez SUV i szereg pokrewnych do niego pomiarów,
3. półautomatycznych metod obliczania objętości mas guzowatych w badaniu PET.

Ocena parametrów anatomicznych w badaniu PET/CT

Zależność pomiędzy wyjściową masą guza (ang. tumour burden) a wynikami leczenia w ziarnicy złośliwej została opisana przeszło 30 lat temu (Schomberg *et al.*, 1984; Specht and Nissen, 1988a). Wyjściowo parametr ten był wyznaczany przez badanie kliniczne, zdjęcie rentgenowskie klatki piersiowej, ewentualnie limfografię, z czasem zastąpione przez tomografię komputerową (Gobbi *et al.*, 2001). Ponieważ wyjściowe metody liczenia całkowitej objętości mas nowotworowych były bardzo czasochłonne, do praktyki klinicznej wszedł szereg zmiennych przybliżających masę guza, takich jak stopień zaawansowania choroby, liczba zajętych węzłów chłonnych, zajęcie narządów pozawęzłowych czy też obecność dużej masy guza (ang. bulk) (Barrington and Meignan, 2019).

Z kolei odpowiedzi na leczenie towarzyszy zmniejszanie się objętości mas nowotworowych. W sposób uproszczony liczenie tego parametru zostało włączone do kryteriów odpowiedzi na leczenie, zwanych kryteriami IWG lub Cheson (Cheson *et al.*, 1999), bazujących na pomiarach prostopadłych wymiarów na przekrojach poprzecznych w TK do sześciu reprezentatywnych zmian. Kolejne uproszczenie tego parametru zaproponowano w kryteriach RECIL - biorących pod uwagę sumę pojedynczych wymiarów jedynie trzech reprezentatywnych zmian (oraz zmianę tego parametru) (Younes *et al.*, 2017). Choć po wprowadzeniu do powszechnego klinicznego użytku badania PET wykorzystanie samych kryteriów tomograficznych do oceny rezultatów

leczenia w ziarnicy złośliwej zostało uznane za dopuszczalne jedynie w przypadku niemożności wykonania badania PET, to jednak niektórzy badacze postulują, że zmniejszanie wielkości i zmiany metaboliczne guza/guzów powinny być brane pod uwagę razem, także w ocenie odpowiedzi na leczenie na wczesnym etapie terapii (Kostakoglu *et al.*, 2012). W pracy tej podkreślane jest znaczenie takiego podejścia szczególnie w sytuacji, jeśli podejrzewa się, że podwyższony wychwyty FDG reprezentuje fałszywie dodatnie wyniki.

Ocena ilościowych parametrów PET - SUV zmiany i jego pochodne

Choć ocena z wykorzystaniem kryteriów Deauville jest zdecydowanie zalecana przez wytyczne dotyczące monitorowania odpowiedzi na leczenie we wszystkich histologicznych podtypach chłoniaków o wysokim wychwycie FDG, to jednak wysoki odsetek wyników fałszywie dodatnich i obawy dotyczące zmienności między obserwatorami wpłynęły na wprowadzenie i rozwijanie narzędzi ilościowych oceny badania PET. Mają one na celu zwiększenie obiektywizacji pomiaru odpowiedzi guza na leczeniu zarówno w codziennej praktyce klinicznej, jak i badaniach naukowych (Kostakoglu and Chauvie, 2018).

Podstawowym sposobem pomiaru metabolizmu ^{18}F -FDG, używanym w półilościowej ocenie badania PET, jest SUV (ang. Standardized Uptake Value – standaryzowana wartość wychwyty), czyli stosunek pomiędzy skorygowanym stężeniem radioznacznika w zadanej jednostce objętości (ang. VOI – Volume Of Interest) do wstrzykniętej pacjentowi dawki w odniesieniu do miary reprezentującej wielkość pacjenta, najczęściej jego masę ciała. W odniesieniu do tej zmiennej powstał szereg jej pochodnych, reprezentujących sposób oznaczania największego wychwyty, takich jak SUVmax, SUVmean czy SUVpeak.

Zmienne te są silnie zależne od szeregu czynników, m. in. typu skanera, parametrów akwizycji czy rekonstrukcji obrazów. Uniemożliwia to m. in. proste porównanie tych parametrów między różnymi ośrodkami. Zaproponowano zatem także

inne sposoby przedstawiania ilościowej aktywności guza. Jednym z nich jest normalizacja wartości SUV względem wartości referencyjnej. W chłoniakach najczęściej przyjmowanym parametrem jest poziom wychwytu wątroby – w zależności od sposobu obliczenia parametr ten określono jako qPET lub rPET (Hasenclever *et al.*, 2014; Annunziata *et al.*, 2016). Innym sposobem wyrażenia tej wielkości jest stosunek SUV zmiany do SUV puli krwi narządów śródpiersia – oznaczone jako SUR. Miara ta została zaproponowana i testowana przede wszystkim przy guzach litych (raku przełyku i niedrobnokomórkowym raku płuca (Bütof *et al.*, 2015, 2018; Shin *et al.*, 2017).

Kolejny wskaźnik ilościowy to $\Delta\text{SUV}_{\text{max}}$, reprezentujący procentową różnicę pomiędzy SUV_{max} w badaniu wyjściowym a wykonywanym w różnych momentach leczenia. Wyniki szeregu badań, szczególnie dotyczących chłoniaków DLBCL dowodzą, że w tych nowotworach zmienna ta jest bardziej powtarzalna i dokładniejsza od analizy wizualnej (Casasnovas *et al.*, 2011; Itti *et al.*, 2013). Przykładowo Rossi *et al.* (2014) na próbie 59 chorych na ziarnicę złośliwą wykazali, że wykorzystanie tej zmiennej przy 71% proggu odcięcia cechuje się znacznie wyższą dodatnią i ujemną wartością predykcyjną niż ocena badania interim w skali 5-stopniowej. Do podobnych wniosków doszli Itti *et al.* (2013), badając populację chorych na DLBCL – zaproponowano próg odcięcia na poziomie 66%.

Metody oceny półautomatycznej - MTV i TLG

Wprowadzenie badania PET sprawiło, że pomiar całkowitej, metabolicznie aktywnej objętości guza został znacząco ułatwiony. Guzy, które gromadzą ^{18}F -FDG, można obrysować w sposób półautomatyczny – przy pomocy ustalonej wartości granicznej SUV - i pomiary ich objętości zsumować, co pozwala na określenie całkowitej metabolicznej objętości guza (MTV). Całkowita glikoliza guza (TLG), jest z kolei pomiarem, w którym bierze się pod uwagę zarówno rozległość zmian, jak i intensywność wychwytu radioznacznika. Wylicza się ją przez przemnożenie MTV przez średnią wartość SUV (Barrington and Meignan, 2019)

Czynnikiem znacząco utrudniającym interpretację i porównanie dotychczas przeprowadzonych badań z wykorzystaniem tych parametrów jest fakt, że obliczenia MTV są prowadzone w bardzo różny sposób przez poszczególne zespoły badawcze. W niedawnym przeglądzie literaturowym (Barrington and Meignan, 2019) wskazali na potrzebę standaryzacji i ujednolicenia sposobu liczenia tych zmiennych.

Główne sposoby wyznaczania wartości granicznej (ang. threshold), będącego bazą dla wyliczania MTV to:

- ustalenie stałego, absolutnego poziomu – np. SUV wynoszącego 2,5 (badania u chorych z wczesną postacią HL - Song *et al.*, 2013; Akhtari *et al.*, 2018; Milgrom *et al.*, 2019) lub 4,0 (wykorzystane w DLBCL - Kurtz *et al.*, 2015),
- SUV w stosunku do punktu odniesienia – np. poziomu wychwytu wątroby lub śródpiersia, jak zaproponowano w kryteriach dla guzów litych – PERCIST (Wahl, Jacene, Kasamon and Martin A Lodge, 2009). W kryteriach tych rekomendacją do wykorzystania przy liczeniu TLG jest wartość będąca pochodną SUV_{mean} wątroby + 3 SD. Podobne podejście w swojej pracy, odnoszącej się do jednego z podtypów chłoniaków (DLBCL), przyjął zespół Ilyas *et al.* (2018).
- stały poziom w stosunku do SUV_{max}, np. wynoszący jego 41% – w zaawansowanej ziarnicy złośliwej takie podejście reprezentują badania odnoszące się do wyjściowego PET (Kanoun *et al.*, 2014), lub wyjściowego badania przed leczeniem nawrotu choroby (Moskowitz *et al.*, 2017). Innym przykładem wykorzystania tego punktu odcięcia może być praca Ben Bouallègue *et al.*, (2017), gdzie z kolei oceniano różne typy chłoniaków z cechą „bulky”.

Zaproponowano też w różnych podtypach chłoniaków bardziej kompleksowe metody przetwarzania obrazów, włączając w to wartości graniczne bazujące na wychwycie na krawędziach zmian, kontury skorygowane w stosunku do tła, metody statystyczne, takie jak analiza skupień (klasteryzacja) czy bazujące na statystyce Bayesowskiej lub Markova (Kostakoglu and Chauvie, 2018; Voltin *et al.*, 2020).

Zdecydowana większość badań dotyczących ziarnicy złośliwej oraz MTV skupia się na ocenie badania wyjściowego (ang. baseline) lub ocenie choroby nawrotowej

(badania wykonywane w określonym momencie – po zakończeniu określonego rodzaju terapii II rzutu lub np. przed ASCT (Moskowitz *et al.*, 2017)).

MATERIAŁY I METODY

Niniejsze badanie jest częścią ostatecznej analizy wyników zindywidualizowanego leczenia chorych z zaawansowaną ziarnicą złośliwą (klinicznie poziom zaawansowania II BX, III i IV) w schemacie rozpoczynającym leczenie I rzutu od chemioterapii eskBEACOPP, w razie dobrej odpowiedzi na pierwsze cykle deeskalowanego do ABVD, realizowanego pomiędzy kwietniem 2003 a sierpniem 2012 w Klinice Hematologii Szpitala Uniwersyteckiego w Krakowie.

Pacjenci byli kierowani do tego sposobu leczenia na podstawie wspólnej decyzji podejmowanej przez lekarza prowadzącego i chorego, po rozmowie na temat możliwych skutków ubocznych leczenia i spodziewanych jego wyników oraz po wyrażeniu świadomej zgody na takie postępowanie. Po wstępnej ocenie wyników leczenia, dokonywanej po pierwszych cyklach (najczęściej dwóch) chemioterapii, dalszą terapię u części osób modyfikowano. W przypadku osób, które dobrze odpowiedziały na leczenie i osiągnęły całkowitą remisję (CR, rozumiana jako Deauville 1-3) terapię deeskalowano do schematu ABVD. Finalnie leczenie pierwszego rzutu składało się z 6 cykli chemioterapii, po których dopuszczano zastosowanie konsolidacyjnej radioterapii (pod postacią IFRT). Końcowa odpowiedź na leczenie była oceniana zgodnie z kryteriami Chesona oraz wyniku badania PET/CT, wykonywanego miesiąc po zakończeniu chemioterapii lub 3 miesiące po konsolidacyjnej radioterapii. Pacjenci z częściową odpowiedzią na leczenie (PR), stabilną chorobą (SD) lub jej progresją (PD) byli kierowani do terapii II rzutu, w przypadkach, w których było to możliwe, uzupełnionej o autologiczny przeszczep komórek macierzystych szpiku (ASCT).

Wszyscy pacjenci otrzymali leczenie wspomagające zgodne z lokalnymi standardami i obejmujące prewencję zespołu lizy guza, w razie wskazań – antybiotykoterapię lub leczenie przeciwwirusowe i przeciwgrzybicze oraz transfuzje preparatów krwiopochodnych (koncentratu krwinek czerwonych lub płytek krwi). Podczas cykli eskBEACOPP stosowano standardowo profilaktykę gorączki neutropenicznej poprzez podawanie czynnika wzrostu kolonii granulocytów (ang. G-CSF – Granulocyte Colony-Stimulating Factor).

Metodyka i dobór badanej grupy

W głównej części niniejszego badania dokonano analizy retrospektywnej historii chorób oraz badań PET/CT wykonywanych u chorych z powyższej grupy, których pełna dokumentacja – w tym obrazy i opisy badań PET/CT wykonywanych na początku leczenia oraz jako wczesna ocena odpowiedzi na leczenie – była dostępna. Ze względu na brak skanera PET/CT w owym czasie w Krakowie, badania te były wykonywane w Zakładzie Medycyny Nuklearnej Centrum Onkologii im. prof. Franciszka Łukaszczyka w Bydgoszczy, skąd zostały udostępnione do celów niniejszej pracy. Badanie ograniczono do okresu od roku 2006 do 2011 ze względu na fakt, że badania obrazowe były wykonywane w jednym ośrodku, według jednego, spójnego protokołu.

kryteria włączenia: pacjenci z zaawansowaną postacią ziarnicy złośliwej (Ann Arbor II BX, III i IV), leczeni w latach 2006 – 2011 w Klinice Hematologii SU w Krakowie początkowo przy pomocy schematu chemioterapii eskBEACOPP, w razie dobrej odpowiedzi klinicznej zmienianego na ABVD, oceniani w badaniu PET/CT przynajmniej w 2 punktach czasowych (przed rozpoczęciem leczenia i jako wczesna ocena skuteczności po 2 cyklu chemioterapii).

kryteria wykluczenia: brak dostępności dokumentacji medycznej i/lub badań PET/CT

Badana populacja została podzielona na dwie grupy: pacjenci z pierwotną opornością lub wznową choroby po leczeniu I rzutu oraz tych, którzy odpowiedzieli na leczenie I rzutu i żyli bez wznowy choroby przynajmniej 5 lat od rozpoznania. W skład badania weszła, poza uwzględnieniem danych klinicznych (takich jak dane demograficzne, stopień zaawansowania choroby, czynniki ryzyka, choroby towarzyszące, szczegóły zastosowanego leczenia, czas wolny od progresji, ewentualny zgon i jego przyczyna), powtórna ocena badań PET/CT wykonywanych przed leczeniem oraz we wczesnym okresie leczenia przez autorkę niniejszej rozprawy – specjalistkę radiologii i diagnostyki obrazowej, nie zaangażowanej wyjściowo w opiekę nad pacjentami. Wykorzystano do niej oprogramowanie komputerowe do przetwarzania badań obrazowych Osirix i Metavol, współpracujących ze standardem DICOM i dostępnych bezpłatnie do celów naukowych.

Badanie uzyskało zgodę Komisji Bioetycznej UJ, o numerze 1072.6120.59.2017.

Wyjściowa ocena badania PET/CT

Pierwszym etapem analizy badań PET/CT była ich podstawowa ocena, wykonywana zgodnie ze standardami z zakresu medycyny nuklearnej i radiologii (uwzględniającymi kryteria Lugano i RECIL), w której zawarto:

- pomiar wartości referencyjnych – MBPS i SUV wątroby,
- rozmieszczenie anatomiczne zmian nowotworowych oraz pomiary ich wielkości,
- SUVmax oraz SUVpeak zmian nowotworowych,
- inne zmiany widoczne w badaniu PET/CT (np. wysięk w opłucnej lub osierdziu oraz inne zmiany chorobowe i warianty anatomiczne, niezwiązane z ziarnicą złośliwą).

Zgodnie z kryteriami Deauville (Meignan *et al.*, 2009), stan chorego w badaniu PET/CT skategoryzowano przy pomocy pięciostopniowej skali oceny (tabela 4):

Tabela 4: Pięciostopniowa skala oceny wychwytu FDG przez zmiany chłoniakowe – zwana skalą Deauville. Punkty na skali odnoszą się do ognisk najwyższego wychwytu w miejscach wyjściowo zajętych przez chorobę (Gallamini and Zwarthoed, 2018).

Punkty	Wartość SUVmax zmiany
1	Nie występuje wzmożony wychwyty
2	Wychwyty FDG \leq MBPS
3	MBPS < wychwyty FDG \leq SUVmax wątroby
4	Wychwyty FDG > wychwyty wątroby
5	Wychwyty FDG \gg wątroby i/lub pojawienie się nowych zmian
X	Pojawienie się nowych ognisk podwyższonego wychwyty, niecharakterystycznych dla chłoniaka

Parametry anatomiczne badania PET/CT

W analizie parametrów anatomicznych bazowano na pomiarach wywodzących się z części tomograficznej badania PET/CT i wzięto pod uwagę trzy rodzaje pomiarów:

- jednowymiarowe, wykonywane na skanach poprzecznych, reprezentowane przez SLD. Zgodnie z kryteriami RECIL, do ich obliczenia brano pod uwagę sumę największych wymiarów na skanach poprzecznych do trzech reprezentatywnych zmian:

$$SLD = \text{wymiarA (zmiana 1)} + \text{wymiarA (zmiana 2)} + \text{wymiarA (zmiana 3)}$$

- dwuwymiarowe, również bazujące na przekrojach poprzecznych – wyrażone przez SPD – sumę iloczynów wymiarów prostokątnych. Zgodnie z zaleceniami kryteriów IWG (Cheson *et al.*, 1999), powtórzonymi także w kryteriach Lugano (Barrington *et al.*, 2014), do ich wyliczenia brano pod uwagę do sześciu reprezentatywnych zmian mierzalnych, z możliwie różnych lokalizacji, większych niż 10x15 mm. Początkowo wyznaczano największy wymiar danej zmiany na przekrojach poprzecznych, następnie na tym samym przekroju – największy prostokąt do nich (ryc. 1). W kolejnych badaniach danego pacjenta pomiary SPD odnoszono do zmian w tych samych lokalizacjach.

$$SPD = \text{wymiar1A} * \text{wymiar1B} + \text{wymiar2A} * \text{wymiar2B} + \dots + \text{wymiar6A} * \text{wymiar6B}$$

- oraz uproszczone wolumetryczne – w tym celu dla każdej ze zmian, uwzględnionych w liczeniu SPD, został zmierzony trzeci wymiar, prostokątny do pierwszych dwóch. Następnie na podstawie trzech pomiarów liczone w uproszczony sposób objętość każdej ze zmian, korzystając ze wzoru dla elipsoidy (dostosowany do faktu, że każdy z wymiarów reprezentował pomiar całości

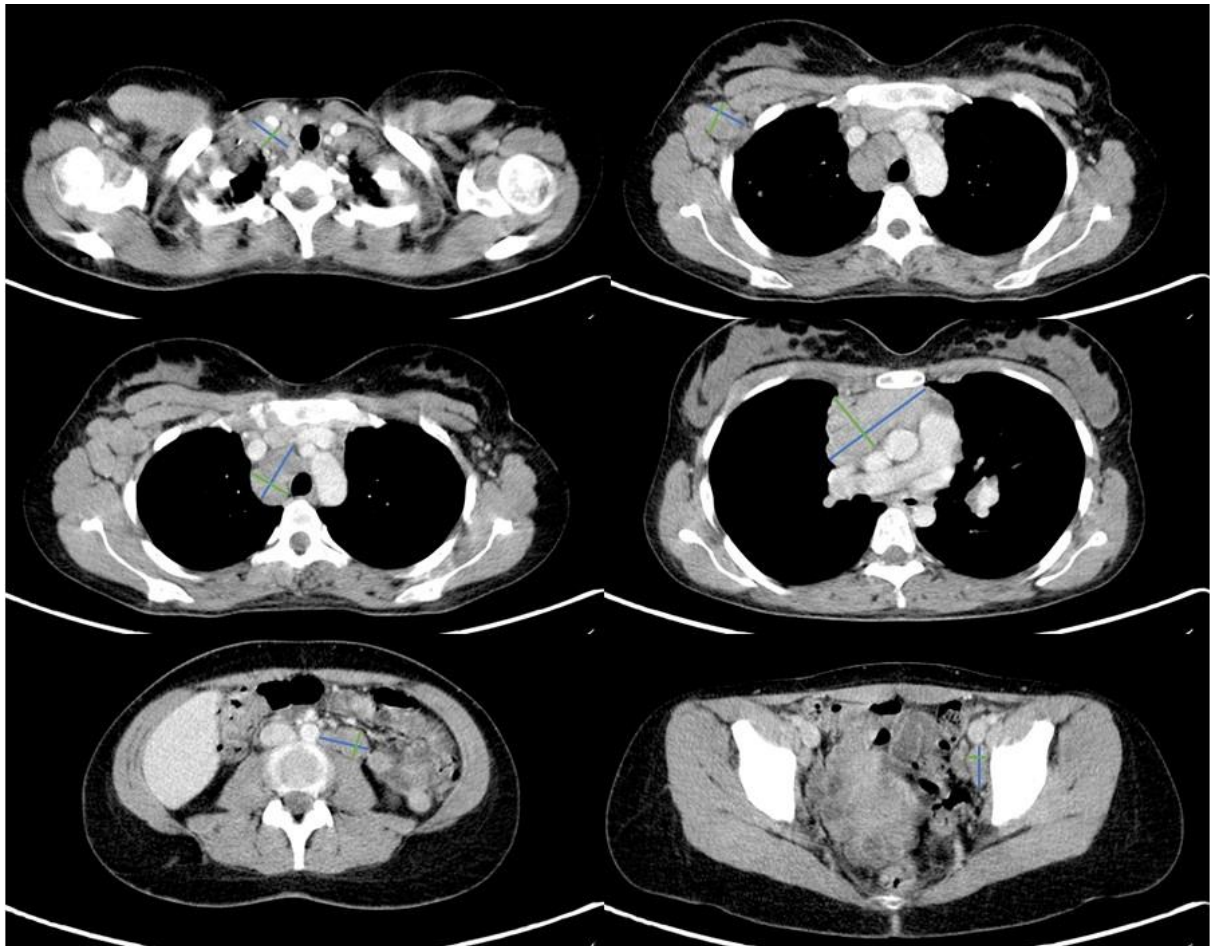
zmiany, a nie długość półosi elipsoidy). Parametrem reprezentującym objętość całkowitą była suma objętości ww. pojedynczych zmian:

$$V_{elipsoidy} = \frac{\pi}{6} * a * b * c$$

$$V_{całkowita} = V_{elipsoidy1} + V_{elipsoidy2} + V_{elipsoidy3} + V_{elipsoidy4} + V_{elipsoidy5} + V_{elipsoidy6}$$

Rycina 1: Sposób wyznaczania SPD – sumy iloczynów najdłuższego wymiaru zmiany na przekroju poprzecznym oraz największego wymiaru prostopadłego do niego, uwzględniająca do sześciu reprezentatywnych zmian mierzalnych.

Część tomograficzna badania PET/CT przed rozpoczęciem leczenia (baseline), pacjentka 7.



Na podstawie badania wyjściowego wybrano zmiany docelowe (ang. target lesions), charakteryzujące się zarówno największą aktywnością metaboliczną, jak i największą wielkością (te dwie cechy zazwyczaj szły w parze) oraz reprezentujące możliwie różne lokalizacje. W ocenie badania interim PET (oraz ewentualnych kolejnych badaniach) odnoszono się do pomiarów w tych samych lokalizacjach oraz ewentualnej zmiany ich wielkości.

Obliczono także procentowe zmiany zmiennych - SLD, SPD i całkowitej objętości wybranych zmian - pomiędzy badaniem wyjściowym (baseline) i wykonywanym wcześniej w cyklu leczenia (interim), zgodnie ze wzorem:

$$\Delta (\text{zmiana zmiennej}) = \frac{\text{zmienna (baseline)} - \text{zmienna (interim)}}{\text{zmienna (baseline)}} * 100\%$$

Badania TK towarzyszące badaniu PET/CT u większości pacjentów wykonywane były bez dożylnego podania kontrastu. Czasami utrudniało to dokładne wyznaczenie granic struktur i narządów. W przypadkach wątpliwych, szczególnie podczas oceny struktur śródpiersia, stosowano wnikliwe porównanie struktur anatomicznych pomiędzy różnymi badaniami u tego samego pacjenta, uwzględniając w tym także – o ile były dostępne - klasyczne badania tomograficzne z okresu diagnozy i leczenia.

Parametry PET

Oceniane parametry obejmowały:

- SUVmax zmian nowotworowych (w wyliczeniu którego brano pod uwagę zmianę charakteryzującą się najwyższym wychwytem FDG w całym organizmie pacjenta i widoczny w jej obrębie punkt cechujący się największą wartością SUV),
- ich SUVpeak, wyznaczany półautomatycznie – jako SUV 4 najgorętszych przylegających do siebie voxelów w obrębie guza,
- SUVmax wątroby,

- SUVmean wątroby - w celu osiągnięcia możliwie największej powtarzalności parametr ten był wyznaczany półautomatycznie na bazie 30 ml VOI przy pomocy programu Metavol, zgodnie z metodyką opisaną przez Hirata i in. (2014),
- wychwyty puli krwi narządów śródpiersia (MBPS, mediastinal blood pool structures) - wyznaczany na podstawie okrągłych ROI (region of interest – obszar zainteresowania) śr. 10 mm, rysowanych na kolejnych przekrojach aorty zstępującej z pominięciem jej ścian i ew. zmian miażdżycowych, charakteryzujących się zwiększonym wychwytem FDG. Do końcowych obliczeń wybierano przekrój, który charakteryzował się największą jednorodnością (reprezentowaną przez najmniejsze odchylenie standardowe SUV).

Na podstawie powyższych zmiennych określano status pacjenta przy wykorzystaniu skali Deauville (jak przy ocenie wyjściowej PET/CT – tabela 4). Wyliczono także kilka dodatkowych zmiennych. Spośród branych pod uwagę parametrów najbardziej obiecujące okazały się: ΔSUV (różnica w SUVmax pomiędzy badaniem wyjściowym – baseline – a wykonywanym wcześniej podczas leczenia – interim) oraz zmienne pozwalające odnieść wychwyty zmiany nowotworowej do wychwyty wątroby (w zależności od sposobu wykonywania obliczeń określane jako rPET i qPET):

$$\Delta SUV = \frac{SUV(baseline) - SUV(interim)}{SUV(baseline)} * 100\%$$

$$rPET = \frac{SUV_{max}}{SUV_{max \text{ wątroby}}}$$

$$qPET = \frac{SUV_{peak}}{SUV_{mean \text{ wątroby}}}$$

Pomiary półautomatyczne

Dodatkowo w analizie PET wykorzystano półautomatyczne metody pozwalające na określenie metabolicznej objętości guza (MTV) – czyli sumy objętości obszarów o wychwyty ¹⁸-FDG wyższym, do zadanego punktu odcięcia - oraz całkowitej glikolizy w

obrębie zmiany (TLG) – iloczynu MTV oraz średniego SUV. Na podstawie danych literaturowych do wyznaczania ww. zmiennych wykorzystano różne punkty odcięcia:

- wartość bezwzględna SUV na poziomie 2,5,
- poziom MBP,
- MBPSmean + SD (odchylenie standardowe, ang. Standard Deviation),
- 41% SUVmax,
- SUVmean wątroby + 3SD (analogicznie do punktu odcięcia przyjętego w kryteriach PERCIST – przy czym przy ocenie retrospektywnej badań PET/CT dysponowano SUV wyliczanym w stosunku do całkowitej masy ciała, a w kryteriach PERCIST zalecono odnoszenie SUV do beztłuszczowej masy ciała) (Wahl, Jacene, Kasamon and Martin A Lodge, 2009).

Z automatycznie wyznaczonych obszarów, charakteryzujących się SUV większym niż punkt odcięcia, wyłączano w sposób półautomatyczny regiony, które fizjologicznie charakteryzują się wysokim nagromadzeniem radioznacznika – np. serce, układy kielichowo-miedniczkowe nerek oraz pęcherz moczowy.

Analiza statystyczna

Zmienne jakościowe (kategoryczne) przedstawiono jako częstość występowania i odpowiadający im procent, natomiast zmienne ilościowe (ciągłe) – jako średnie \pm odchylenie standardowe (SD) lub mediany i rozstęp międzykwartyłowy (ang. IR – Interquartile Range), w zależności od normalności rozkładu. Normalność rozkładu oszacowano przy pomocy testu Shapiro-Wilka. Różnice w rozkładzie zmiennych ciągłych pomiędzy grupami pacjentów osiągających długie przeżycie wolne od choroby w stosunku do tych z progresją lub wznową choroby zostały określone przy pomocy dwustronnego testu U Manna-Whitney'a lub testu t-Studenta. Ze względu na niewielką liczbę przypadków, dla oceny związku pomiędzy zmiennymi kategorycznymi a statusem zdrowotnym pacjentów (pięcioletnie przeżycie wolne od choroby po leczeniu I rzutu vs. wznowa lub progresja) wykorzystano dokładny test Fishera.

Model jednoczynnikowej regresji logistycznej został zbudowany z wykorzystaniem statusu zdrowotnego pacjenta (pacjenci z pierwotną opornością lub wznową choroby po leczeniu I rzutu vs. ci, którzy odpowiedzieli na leczenie I rzutu i żyli bez wznowy choroby przynajmniej 5 lat od rozpoznania) jako zmiennej zależnej oraz każdej z badanych zmiennych jako predyktora. Dla każdego z modeli sprawdzono obecność punktów odstających. Jednym ze wskaźników, służących do oceny różnych modeli i wyboru najlepszych predyktorów było kryterium informacyjne Akaike (ang. AIC - Akaike Information Criterion), odnoszące się do dobroci dopasowania modelu. Ze względu na wysoką korelację pomiędzy wszystkimi badanymi predyktorami nie dopasowywano modelu wielowymiarowej regresji logistycznej. Dla rokujących zmiennych ilościowych wykreślono krzywe ROC (ang. Receiver Operator Curve), a następnie wyznaczono punkt odcięcia, za pomocą indeksu Youdena lub metody stycznej.

Obliczenia statystyczne zostały wykonane przy wykorzystaniu programu STATISTICA v. 13.3 (na licencji zakupionej przez UJ CM). We wszystkich zastosowanych testach przyjęto 5% poziom istotności statystycznej.

WYNIKI

Pomiędzy rokiem 2003 a 2012 w Klinice Hematologii UJCM łącznie 188 pacjentów z zaawansowaną ziarnicą było leczonych przy pomocy schematów leczenia rozpoczynającego się od eskBEACOPP. W grupie tej 161 pacjentów (85,5%) po leczeniu I rzutu uzyskało całkowitą remisję, natomiast czterech (2,1%) częściową. U pozostałych 23 osób (12,2%) stwierdzono pierwotną oporność na leczenie. Dwie osoby zmarły podczas leczenia, z przyczyn niezwiązanych bezpośrednio z progresją choroby (jedna osoba z powodu zakrzepicy pnia mózgu, druga – na skutek powikłań po zapaleniu trzustki). W okresie obserwacji, którego mediana wyniosła 10,2 lat, zmarło kolejne 21 osób – w tym 20 z powodu progresji lub wznowy choroby, a jedna w wypadku drogowym.

Deeskalacja była możliwa u 141 osób (75%). Uzupełniającą radioterapię zastosowano u 64 osób. U osób, których odpowiedź na leczenie I rzutu była niewystarczająca, zastosowano leczenie II rzutu, u 14 osób, które na niego odpowiedziały, skonsolidowanego przez ASCT.

W dziesięcioletnim okresie obserwacji przeżywalność ogólna (OS) w całej populacji wyniosła 89,4%, natomiast przeżycie bez progresji (PFS) osiągnęło 79,3% chorych. W grupie chorych, którzy byli leczeni sześcioma cyklami eskBEACOPP znamienne częściej występowały hematologiczne i zatorowe powikłania, włączając w to gorączkę neutropeniczną (25 [53,2%] pacjentów vs. 7 [5%]), ciężką anemię (35 [74,5%] vs. 11 [7,8%]), trombocytopenię (16 [34%] vs. 7 [5%]) oraz zatorowość płucną (3 [6,4%] vs. 0) w stosunku do chorych z deeskalacją terapii.

Opis badanej próby

Do finalnej analizy włączono 60 pacjentów z zaawansowaną ziarnicą złośliwą. Mediana okresu obserwacji wyniosła 10,2 lat. W tym czasie pięciu pacjentów zmarło, z czego trzech na skutek progresji choroby Hodgkina. Pozostała dwójka pacjentów to

wyżej opisani, którzy zmarli podczas leczenia I rzutu bez ewidentnych cech progresji choroby. Zostali oni wykluczeni z końcowej analizy, ponieważ nie można było w sposób wiarygodny ocenić ich odpowiedzi na leczenie. Pośród osób, które przeżyły do końca okresu obserwacji, były trzy przypadki oporności na leczenie oraz dwa przypadki wznowy choroby po zakończeniu leczenia. Wszyscy oni zostali poddani leczeniu drugiego rzutu, w trzech przypadkach uzupełnionego o autologiczny przeszczep komórek macierzystych (ASCT).

W tabeli 5 zawarto charakterystykę kliniczną całej populacji poddanej leczeniu rozpoczynającemu się od schematu eskBEACOPP w Klinice Hematologii SU w Krakowie w porównaniu do pacjentów włączonych do finalnej analizy. Płeć, skale zaawansowania i ryzyka, obecność objawów ogólnych, deeskalację, radioterapię i deeskalację wyrażono jako liczbę pacjentów i ich odsetek w danej grupie. Ze względu na brak rozkładu normalnego wiek jest zaprezentowany jako mediana (rozstęp międzykwartylowy). Mimo istnienia potencjalnych przyczyn dysproporcji między grupami (np. ze względu na niedostępność skanera PET w tym czasie w Krakowie i znikomą liczbę takich aparatów w Polsce, badania PET/CT mieli wykonywane jedynie ci chorzy, którzy byli w na tyle dobrym stanie ogólnym, by sprostać drodze tam i z powrotem do Zakładu Medycyny Nuklearnej Centrum Onkologii w Bydgoszczy - czyli blisko 1000-kilometrowej podróży), obie grupy są bardzo zbliżone pod względem procentowego rozkładu częstości występowania określonych podtypów histologicznych ziarnicy złośliwej, obecności czynników ryzyka oraz rodzaju leczenia (w tym deeskalacji terapii, konsolidacyjnej radioterapii czy ASCT). Różnią się one natomiast proporcją mężczyzn i kobiet – w grupie poddanej finalnej analizie znalazło się więcej kobiet (42% vs. 62,1%).

Tabela 5: Wyjściowa charakterystyka pacjentów leczonych w Klinice Hematologii SU w Krakowie wg. schematu eskBEACOPP oraz podgrupy włączonej do finalnego etapu badania.

Parametr	Wszyscy pacjenci (n = 188)	Pacjenci włączeni do finalnej analizy (n = 58)
Płeć (mężczyźni)	109 (58%)	22 (37,9 %)
Wiek	33 (18-59)	29 (23-36)
Ann Arbor		
II BX	69 (36,7%)	16 (27,6%)
III	45 (23,9%)	14 (24,1%)
IV	74 (39,4%)	28 (48,3%)
Objawy ogólne	139 (73,9%)	38 (65,5%)
IPS		
0-2	88 (46,8%)	27 (46,6%)
3-7	100 (53,2%)	31 (53,4%)
Deeskalacja	45 (77,6%)	45 (77,59%)
Radioterapia	64 (34%)	21 (36,21%)
Konsolidacja ASCT	14 (7,5%)	4 (6,9%)

Grupę włączoną do finalnej analizy podzielono na dwie podgrupy – pacjentów z pierwotną opornością lub wznową choroby po leczeniu I rzutu oraz tych, którzy odpowiedzieli na leczenie I rzutu i żyli bez wznowy choroby przynajmniej 5 lat od rozpoznania. Szczegółowe przedstawienie rozkładu powyższych cech w zależności od podgrupy przedstawiono w tabeli 6. Podobnie, jak w poprzedniej tabeli płeć, skale zaawansowania i ryzyka, obecność objawów ogólnych, deeskalację, radioterapię i deeskalację wyrażono jako liczbę pacjentów i ich odsetek w danej grupie. Ze względu na brak rozkładu normalnego wiek jest zaprezentowany jako mediana (rozstęp międzykwartylowy). Przed leczeniem nie stwierdzono statystycznie istotnych różnic pomiędzy pacjentami, którzy pozostali wolni od choroby w przeciągu 5 lat od leczenia I rzutu a tymi, u których wystąpiła wznowa lub progresja. Natomiast czynnikami, które statystycznie istotnie różnią się w obu grupach, jest odsetek chorych, u których podjęto decyzję o deeskalacji leczenia oraz konsolidacji terapii przez wysokodawkową chemioterapię wspomaganą przeszczepieniem komórek macierzystych (ASCT).

Tabela 6: Wyjściowa charakterystyka pacjentów włączonych do badania.

Parametr	Wszyscy pacjenci (n = 58)	Brak wznowy lub progresji (n = 50)	Progresja lub wznowa choroby (n = 8)	Wartość p
Płeć (mężczyźni)	22 (37,9 %)	13 (34 %)	5 (62,5 %)	0,09
Wiek	29 (23-36)	30 (24-36)	26 (23-33)	0,43
Ann Arbor				
II	16 (27,6%)	15 (30%)	1 (12,5%)	0,42
BX	14 (24,1%)	12 (24%)	2 (25%)	1,00
III	28 (48,3%)	23 (46%)	5 (62,5%)	0,46
IV				
Objawy ogólne	38 (65,5%)	33 (66%)	5 (62,5%)	1,00
IPS				
0	6 (10,4%)	6 (12%)	0 (0%)	0,58
1	8 (13,8%)	8 (16%)	0 (0%)	0,58
2	13 (22,4%)	12 (24%)	1 (12,5%)	0,67
3	13 (22,4%)	9 (18%)	4 (50%)	0,07
4	11 (19%)	9 (18%)	2 (25%)	0,64
5	7 (12,1%)	6 (12%)	1 (12,5%)	1,00
Deeskalacja	45 (77,6%)	42 (84%)	3 (37,5%)	0,01
Radioterapia	21 (36,2%)	18 (36%)	3 (37,5%)	1,00
Konsolidacja ASCT	4 (6,9%)	1 (2%)	3 (37,5%)	0,01

Leczenie chorych rozpoczynano od schematu eskBEACOPP. Po wstępnej ocenie wyników leczenia dalszą terapię u części osób modyfikowano, zgodnie z przyjętym protokołem instytucjonalnym. W przypadku osób, które dobrze odpowiedziały na leczenie, osiągając całkowitą remisję w badaniu interim PET (Deauville 1-3), chemioterapię deeskalowano do schematu ABVD (w finalnej grupie w 6 przypadkach po 2 cyklach eskBEACOPP, w 36 – po 3 cyklach eskBEACOPP, w 8 – po 4 cyklach). W całości leczenie pierwszego rzutu składało się z 6 cykli chemioterapii. W grupie poddanej szczegółowej analizie badań obrazowych w 21 przypadkach wyjściowe leczenie skonsolidowano przez radioterapię (pod postacią IFRT), u 4 osób – przez autologiczny przeszczep komórek macierzystych (ASCT).

U wszystkich pacjentów, włączonych do niniejszego badania, obrazowanie przy pomocy PET/CT było wykonywane co najmniej dwukrotnie podczas leczenia: jako wyjściowa ocena stopnia zaawansowania choroby (ang. staging) i jako badanie wczesnej odpowiedzi na zastosowane leczenie (ang. interim). U większości chorych (46 osób) było

dostępne także badania po zakończeniu terapii I rzutu (zgodnie z wytycznymi wykonywane zazwyczaj 3 tygodnie po chemioterapii lub trzy miesiące po IFRT). Badanie interim u większości pacjentów zostało wykonane po drugim cyklu chemioterapii (z wyjątkiem 9 pacjentów, u których wykonano je po 3 cyklu). Nie stwierdzono statystycznie istotnej różnicy w badanych parametrach w zależności od umiejscowienia w czasie badania kontrolnego. Ze względu na odpowiedź różnych tkanek, głównie szpiku i zmian nowotworowych, na chemioterapię, badanie interim wykonywano tak blisko kolejnego cyklu leczenia, jak tylko było to możliwe – zazwyczaj w 8 tygodniu od rozpoczęcia leczenia.

Dane pozyskane z badania PET/CT, wykonanego jako ocena wczesnej odpowiedzi na leczenie (interim), zostały zanalizowane w całej badanej populacji oraz – osobno – w podgrupie, która osiągnęła remisję po leczeniu I rzutu trwającą przynajmniej 5 lat i u pacjentów z progresją lub wznową choroby

Wyjściowa ocena badania PET/CT

U wszystkich chorych w badanej populacji choroba wyjściowo manifestowała się przez powiększenie węzłów chłonnych, cechujących się metabolizmem FDG glukozy znacznie wyższym od SUV wątroby. Przy czym u wszystkich chorych były zajęte węzły chłonne z piętér ponad przeponą (szczególnie śródpiersia i wnęk – 55 pacjentów oraz szyjne i nadobojczykowe – u 47 chorych). Węzły chłonne w lokalizacjach podprzeponowych były zajęte u 42 pacjentów.

Zajęciu węzłów chłonnych towarzyszyły inne widoczne zmiany:

- płyn w opłucnej – obecny u 6 pacjentów,
- niedodma – u 2 pacjentów,
- płyn w osierdziu – u 5 osób,
- widoczne badaniu TK w topografii nadnerczy guzy o patologicznym wychwyte FDG – u 2 pacjentów,
- ognisko w wątrobie, wyodrębniające się zarówno w badaniu PET, jak i TK (1 pacjent),
- ogniskowy wzrost wychwytu FDG w tarczycy (2 pacjentów).

Oceniając część tomograficzną badania PET odkryto też zmiany, które nie są charakterystyczne dla ziarnicy złośliwej i nie zmieniały się w trakcie terapii (najpewniej ich etiologia była inna). Były to: zmiana torbielowata w przymaciczu, torbiel śledziony, zmiana lito-tłuszczowa w obrębie mięśnia biodrowo-lędźwiowego, naczyniak trzonu kręgu i pojedyncze stacjonarne guzki (w tym także uwapnione) w obrębie mięszu płucnego.

Niemal u wszystkich pacjentów (56 osób) stwierdzono w badaniu przed rozpoczęciem leczenia poziom metabolizmu FDG struktur kostnych wyższy od poziomu SUV wątroby. Przy czym w większości przypadków było to rozlane podwyższenie SUV, o wartościach od 2,6 do 6,4. U 13 pacjentów stwierdzono ogniskowy wzrost SUV szpiku kostnego, natomiast tylko u 2 osób nieprawidłowemu obrazowi struktur kostnych w PET towarzyszyły zmiany widoczne w badaniu TK (pod postacią ognisk sklerotycznych).

Ocena parametrów anatomicznych badania PET/CT

W pierwszej kolejności przedstawione zostaną wyniki odnoszące się do parametrów anatomicznych badania PET/CT, wraz z odniesieniem do złotego standardu oceny badania, jakim jest ocena w skali Deauville (tabela 7). W tabeli tej wartości są zaprezentowane dla skali Deauville jako częstość występowania i odpowiadający im procent w obrębie danej grupy, dla innych wartości - jako średnia (odchylenie standardowe) lub mediana (rozstęp międzykwartylowy) przy braku rozkładu normalnego.

Tabela 7: Wyniki analizy danych z badania PET/CT, wykonanego jako ocena wczesnej odpowiedzi na leczenie.

Parametr	Wszyscy pacjenci (n = 58)	Brak wznowy lub progresji (n = 50)	Progresja lub wznowa choroby (n = 8)	Wartość p
Deauville (interim)				
1	1 (1,7 %)	1 (2,0 %)	0 (0,0 %)	1,00
2	17 (31,0 %)	17 (34,0 %)	0 (0,0 %)	0,090
3	26 (44,8 %)	25 (50,0 %)	1 (12,5 %)	0,063
4	12 (20,7 %)	7 (14,0 %)	5 (62,5 %)	0,007
5	2 (3,5 %)	0 (0,0 %)	2 (25,0 %)	0,0167
SLD_{wyjściowe} [cm]	15,03 (3,8)	14,57 (3,71)	17,89 (3,16)	0,020
SLD_{interim} [cm]	8,6 (3,4)	8,03 (3,07)	12,15 (3,37)	0,001
ΔSLD [%]	40,57 (31,69-52,5)	41,7 (33,0-53,49)	31,54 (26,11-42,23)	0,044
SPD_{wyjściowe} [cm²]	62,23 (45,91-90,56)	57,56 (42,72-87,98)	92,72 (60,92-133,47)	0,003
SPD_{interim} [cm²]	20,97 (12,02-35,12)	17,49 (11,61-31,99)	39,55 (33,51-72,79)	0,008
ΔSPD [%]	66,93 (14,59)	68,78 (13,15)	55,43 (18,63)	0,007
V_{wyjściowe} [cm³]	224,415 (116,541-404,793)	175,242 (112,393-364,85)	520,042 (305,916-600,97)	0,014
V_{interim} [cm³]	39,804 (14,967-84,568)	27,019 (13,708-77,572)	148,267 (71,287-312,45)	0,001
ΔV [%]	80,44 (71,40-90,67)	85,03 (72,69- 91,64)	70,45 (56,86-78,39)	0,014

W chwili diagnozy u wszystkich pacjentów zmiany chłoniakowe charakteryzowały się wartością SUV znacznie wyższą niż SUV wątroby, dlatego też zostali oni sklasyfikowani na najwyższym, piątym poziomie skali Deauville. W przypadku wczesnej oceny w badaniu PET/CT u 18 pacjentów SUVmax zmian chorobowych był niższy niż MBPS (punkty 1 i 2 na skali pięciostopniowej Deauville), u kolejnych 26 – wyższy niż MBPS, lecz niższy niż SUV wątroby (Deauville 3). Jedynie w 14 przypadkach choroba była aktywna zgodnie z wytycznymi dotyczącymi analizy SUV (Deauville 4-5). W podgrupach pacjentów z lub bez progresji/wznowy choroby, odsetki pacjentów, których SUV zmian nowotworowych został zaklasyfikowany do grupy 4 lub 5, różnią się znamienne – przy poziomie wychwytu radioznacznika przez zmiany nowotworowe wyższego, niż poziom referencyjny wątroby, są wyższe w grupie, w której leczenie I rzutu było nieskuteczne. Istotną statystycznie różnicę stwierdzono także dla wszystkich przebadanych parametrów, wiążących się z oceną anatomiczną. W ocenie SLD, SPD i analizie objętościowej, zarówno dla badań wyjściowych, jak i wykonywanych wcześniej w trakcie leczenia, otrzymano wyższe wartości u pacjentów z progresją lub wznową choroby, w porównaniu do grupy z remisją trwającą przynajmniej 5 lat po leczeniu I rzutu. Odwrotnie zachowywały się zmiany ww. parametrów – zarówno dla SLD, SPD, jak i wolumetrii były mniej zaakcentowane u pacjentów, u których leczenie I rzutu okazało się nieskuteczne. Wyniki te potwierdzają tezę, że wyżej wymienione parametry anatomiczne mogą być użyteczne w ocenie odpowiedzi na leczenie w ziarnicy złośliwej.

W celu dokładniejszej oceny znaczenia SLD, SPD i obliczeń wolumetrycznych dla określenia rokowania w ziarnicy złośliwej, zbudowano model jednoczynnikowej regresji logistycznej, w którym status pacjenta określono jako zmienną zależną, natomiast testowane parametry – jako zmienne niezależne. Wyniki zostały podsumowane w tabeli 8, gdzie podano także wartość ilorazu szans (ang. OR – Odds Ratio) oraz dodatkowy parametr Akaike dobroci dopasowania (AIC). W tabeli nie uwzględniono wyników regresji logistycznej jednoczynnikowej dla płci, wyjściowego zaawansowania choroby (IPS) oraz radioterapii, ponieważ nie były one istotne statystycznie (dla modelu uwzględniającego płeć wartość p wynosiła 0,28, dla IPS – 0,22, dla radioterapii – 0,7). W przyjętym modelu najniższe wartości AIC zostały osiągnięte dla objętości zmian w badaniu wyjściowym oraz parametrów mierzonych w badaniu PET/CT wykonywanym

jako wczesna ocena wyników leczenia (interim) – wyników w skali Deauville, objętości, SPD i SLD. Zmiana objętości, podobnie jak pomiary SPD w badaniu przed rozpoczęciem leczenia, charakteryzowały się pośrednimi wartościami kryterium dobroci dopasowania Akaike. Najwyższe wartości AIC charakteryzowały zmianę SLD, SPD oraz pomiar SLD w badaniu przed rozpoczęciem leczenia.

Tabela 8: Wyniki jednoczynnikowej regresji logistycznej oraz miara dobroci dopasowania Akaike (AIC).

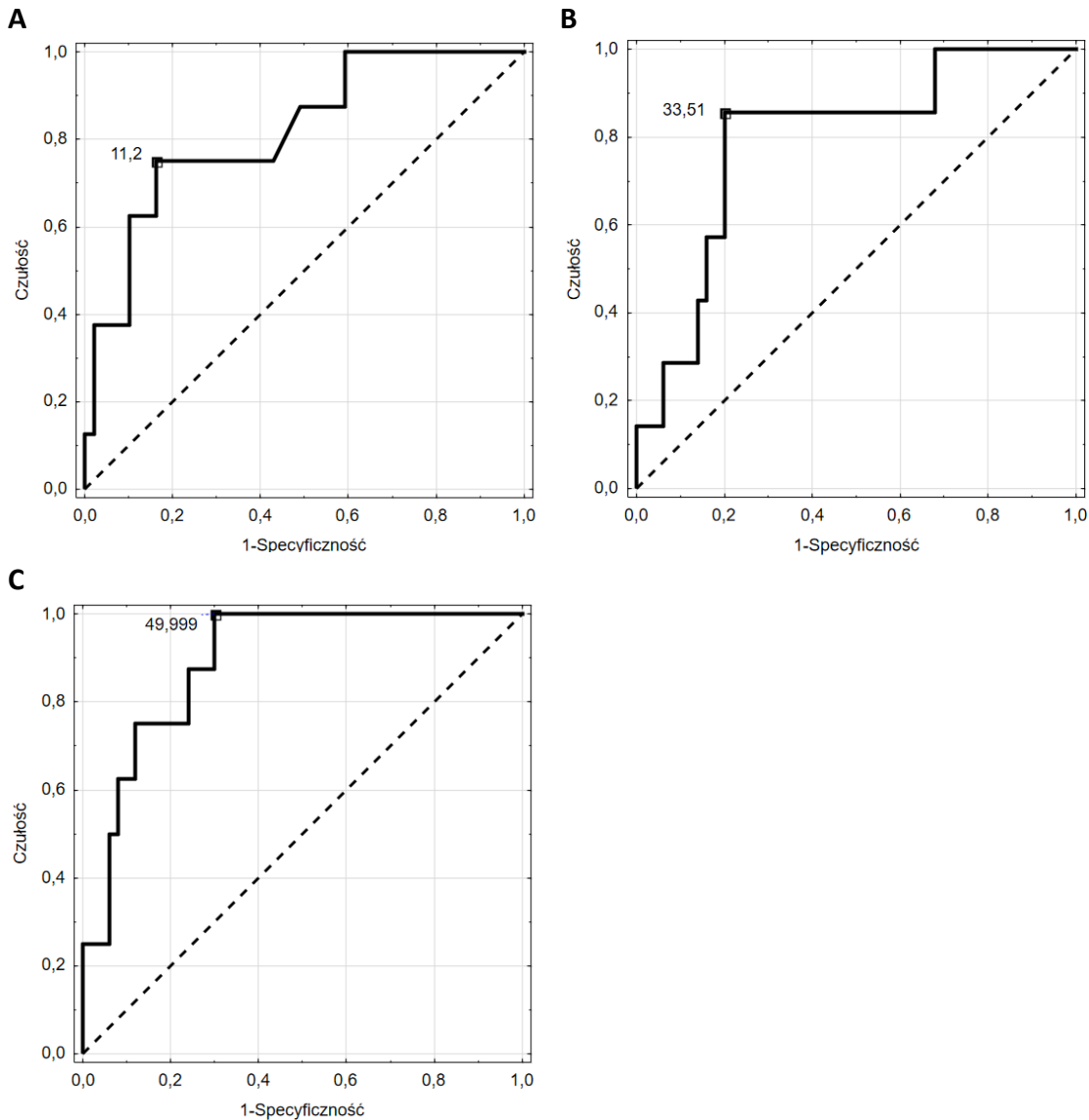
Parametr	OR (95% CI)	Wartość p	AIC
Skala Deauville(interim) [*10]	1,34 (1,09-1,65)	0,01	33,9
SLD_{wyjściowe}	1,27 (1,02-1,58)	0,03	45,23
SLD_{interim}*	1,48 (1,13-1,94)	0,005	40,18
ΔSLD* [/10]	2,05 (1,02-4,1)	0,043	45,09
SPD_{wyjściowe}* [/10]	1,3 (1,02-1,66)	0,034	41,3
SPD_{interim}* [/10]	1,76 (1,1-2,8)	0,018	40,2
ΔSPD [/10]	1,89 (1,08-3,32)	0,027	44,8
V_{wyjściowe}* [/100]	1,56 (1,01-2,42)	0,045	37,7
V_{interim}* [/100]	4,66 (1,73-12,54)	0,002	36,9
ΔV* [/10]	1,97 (1,03-3,8)	0,041	41,9

* Wartości odstające zostały wyłączone z analizy.

Dla najbardziej obiecujących parametrów (SLD_{interim}, SPD_{interim} oraz V_{interim}) zostały wykreślone krzywe ROC (Receiver Operating Characteristic), które posłużyły do wyznaczenia punktów odcięcia przy pomocy metody Newtona (ryc. 2). W przypadku SLD_{interim} (ryc. 1A) ta wartość wyniosła 11,2 cm, dla SPD_{interim} (ryc. 1B) – 33,51 cm², natomiast dla objętości V_{interim} (ryc. 1C) – prawie 50 cm³ (49,999 cm³). Wszystkie wykreślone krzywe ROC były istotne statystycznie. Dla zbadanych zmiennych pole pod krzywą (Area Under the Curve – AUC) było największe dla V_{interim} i wyniosło 0,874 (CI 95%: 0,766-0,981), z kolei dla SLD_{interim} i SPD_{interim} jego wartość była zbliżona i wynosiła odpowiednio 0,821 (CI 95%: 0,667-0,975) i 0.818 (CI 95%: 0,643-0,993). Różnice

między dowolnymi dwiema parami pól pod wykreślonymi krzywymi nie były istotne statystycznie, co sugeruje, że każdy z przedstawionych parametrów może zostać użyty dla oceny postępów leczenia w ziarnicy złośliwej.

Rycina 2: Krzywe ROC dla $SLD_{interim}$ (A), $SPD_{interim}$ (B) i $V_{interim}$ (C) oraz proponowane punkty odcięcia.



Dla ww. parametrów - $SLD_{interim}$, $SPD_{interim}$ i $V_{interim}$ - zostały policzone czułość, swoistość, oraz dodatnia i ujemna wartość predykcyjna (PPV i NPV) (tabela 9). Te same wskaźniki zostały wyliczone dla klasyfikacji aktywności metabolicznej zmian chłoniakowych w badaniu interim PET/CT, z wykorzystaniem zarówno IHP (skala 5-cio

stopniowa Deauville 1-2 vs 3-5) oraz kryteriów Lugano (punktem odcięcia jest poziom wychwytu wątroby). Tabela została uzupełniona przez klasyfikację pacjentów bazującą na połączeniu pomiarów anatomicznych i metabolicznych. Zastosowano wartości punktów odcięcia bazujące na analizie powyżej zaprezentowanych krzywych ROC.

Tabela 9: Podsumowanie wyników różnych metod interpretacji badania interim PET/CT – bazujących na skali 5-cio stopniowej Deauville oraz pomiarów anatomicznych, wywodzących się z części tomograficznej.

	IHP (punkt odcięcia MBP) [%]	Lugano (punkt odcięcia SUV wątroby) [%]	SLD_{interim} (punkt odcięcia 11,2 cm) [%]	SPD_{interim} (punkt odcięcia 33,51 cm ²) [%]	V_{interim} (punkt odcięcia 49,99 cm ³) [%]	łącznie Lugano i SLD_{interim} (punkt odcięcia 11,2 cm) [%]	łącznie Lugano i SPD_{interim} (punkt odcięcia 33,51 cm ²) [%]	łącznie Lugano i V_{interim} (punkt odcięcia 49,99 cm ³) [%]
Czułość	100	87,5	75	87,5	100	50	75	87,5
Swoistość	36	86	84	80	68	100	98	94
PPV	20	50	42,9	41,2	33,3	100	86	70
NPV	100	97,7	95,5	97,6	100	92,6	96,1	97,9

Zastosowanie kryteriów oceny bazujących na części PET badania – zarówno bazujących na kryteriach IHP, jak i Lugano – pozwoliło na osiągnięcie bardzo wysokiej ujemnej wartości predykcyjnej – odpowiednio na poziomie 100% i 97,7%. Jednakże taki sposób oceny badania interim wiązał się z niską dodatnią wartością predykcyjną tego badania (odpowiednio 20% i 50%). Zbliżona sytuacja miała miejsce, jeśli do oceny użyto wyłącznie wskaźników tomograficznych. Połączenie kryteriów oceny części nuklearnej i anatomicznej badania PET/CT - uznanie za dodatnie badań spełniających kryterium Lugano oraz o dużej powierzchni lub objętości zmian (odpowiednio przekraczającej dla sumy sześciu najbardziej reprezentatywnych zmian 33,51 cm² lub 49,99 cm³) pozwoliło na uzyskanie znacznie wyższych dodatnich wartości predykcyjnych, przy zachowaniu wysokiej czułości i swoistości.

Ocena parametrów związanych z SUV

W tabeli 10 podsumowano statystyki opisowe badanych parametrów oraz różnice pomiędzy grupami pacjentów z progresją lub wznową choroby oraz bez cech choroby w 5 lat po zakończeniu leczenia I rzutu. Statystycznie istotne różnice pomiędzy badanymi grupami stwierdza się w przypadku czterech zmiennych:

- SUVmax interim, którego wartości są znacznie niższe dla grupy o korzystnym wyniku leczenia I rzutu – z medianą na poziomie odpowiednio 2,36 vs. 4,25,
- Stosunku pomiędzy SUV docelowych zmian nowotworowych i wątroby, wyrażonego zarówno w postaci SUVmax, jak i SUVpeak – qPET lub rPET, z wartościami znacznie wyższymi w grupie z progresją lub wznową choroby (1,44 vs. 0,78 oraz 1,59 vs. 0,84) w stosunku do grupy pacjentów z dobrą odpowiedzią na leczenie I rzutu
- Zmiany SUVmax pomiędzy badaniem interim a wyjściowym (Δ SUVmax), którego wartość jest wyższa dla bez wznowy lub progresji choroby niż w grupie o niekorzystnym wyniku leczenia I rzutu – mediana 84,13 (rozstęp międzykwartylowy 81,70-87,44) vs. 73,77 (58,76-79,94).

Tabela 10: Wyniki analizy parametrów badania PET.

Parametr	Wszyscy pacjenci (n = 58)	Brak progresji lub wznowy (n = 50)	Progresja lub wznowa choroby (n = 8)	Wartość p
SUVmax (wyjściowe)	15,3 (12,33-18,08)	15,25 (12,27-17,43)	16,73 (13,06-19,6)	0,572
SUVpeak (wyjściowe)	12,69 (9,75-14,37)	12,66 (9,72-14,42)	13,14 (10,62-13,88)	0,814
rPET (wyjściowe)	5,49 (4,19-7,24)	5,36 (4,19-6,98)	6,42 (4,43-7,32)	0,528
qPET (wyjściowe)	5,74 (4,88-7,96)	5,72 (4,91-7,81)	7,64 (4,88-8,45)	0,435
SUVmax (interim)	2,44 (1,95-2,93)	2,36 (1,82-2,72)	4,25 (3,57-4,72)	<0,001
SUVpeak (interim)	2,07 (1,61-2,40)	2,04 (1,57-2,22)	2,87 (2,07-3,51)	0,269
rPET (interim)	0,80 (0,67-0,96)	0,78 (0,67-0,92)	1,44 (1,24-1,93)	<0,001

Parametr	Wszyscy pacjenci (n = 58)	Brak progresji lub wznowy (n = 50)	Progresja lub wznowa choroby (n = 8)	Wartość p
qPET (interim)	0,91 (0,7-1,14)	0,84 (0,69-1,04)	1,59 (0,99-2,15)	0,002
Δ SUVmax	83,86 (77,63-87,29)	84,13 (81,70-87,44)	73,77 (58,76-79,94)	0,004

Z uwagi na brak normalności rozkładów przedstawione zostały mediany zmiennych (rozstęp międzykwartyłowy).

Dla najbardziej obiecujących zmiennych (rPET, qPET, SUVmax (interim) i Δ SUVmax) obliczono wyniki regresji jednoczynnikowej. Rezultaty zostały podsumowane w tabeli 11, porównano je także do „złotego standardu” – wartości SUVmax zaprezentowanej w skali 5-cio stopniowej Deauville (tabela 8). W przypadku tej analizy wartość SUVmax (interim) i skala Deauville okazały się mieć najniższe wartości AIC (przy bardzo podobnych wartościach ilorazu szans). Wartości p modelu dla rPET, qPET i Δ SUV okazały się być bardzo zbliżone – przy czym ostatnia z tych zmiennych okazała się mieć wyższe AIC oraz szerszy przedział ufności ilorazu szans. Spośród ww. zmiennych o AIC mniejszym od 35, rPET charakteryzuje się najwyższą liczbą pacjentów poprawnie zaklasyfikowanych do grupy, która doświadczyła progresji lub wznowy choroby (por. tabela 16), z kolei skala Deauville badania interim okazała się lepsza w determinowaniu przynależności pacjenta do grupy bez cech choroby w 5 lat po zakończeniu leczenia I rzutu.

Tabela 11: Wyniki modeli jednoczynnikowej analizy regresji logistycznej dla najbardziej obiecujących zmiennych z badania PET.

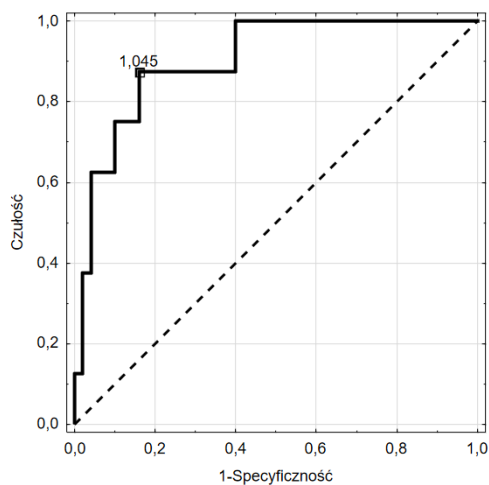
Zmienne	OR (95% CI)	Wartość p	AIC
rPET (interim)*[*10]	1,54 (1,03-2,3)	0,035	31,3
qPET (interim)*[*10]	1,69 (1,01-2,82)	0,045	26,3
SUVmax (interim)*[*10]	1,34 (1,09-1,65)	0,005	22,9
Δ SUV [/10]	3,63 (1,1-12)	0,034	41

* Wartości odstające zostały wyłączone z analizy.

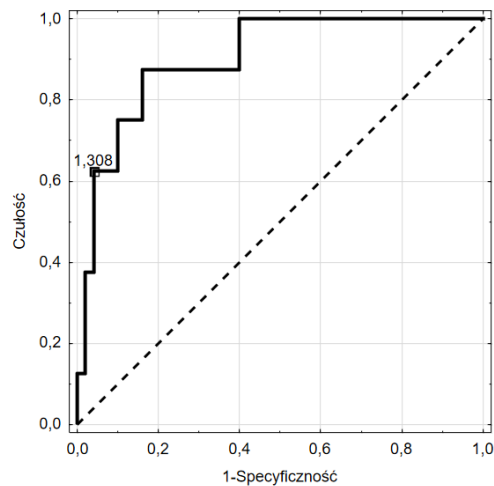
W kolejnym kroku opracowano krzywe ROC dla wartości rPET i qPET w badaniu interim. Nie stwierdzono istotnych statystycznie różnic w zakresie pól pod krzywą tych parametrów, niemniej pole pod krzywą (AUC) miało nieco wyższą wartość dla rPET_{interim} niż qPET_{interim}: odpowiednio 0,903 (CI: 0,802-1) vs. 0,851 (CI: 0,725-0,978). Zaproponowano próg odcięcia dla parametru rPET na poziomie 1,05 lub 1,31, w zależności od przyjętej metody jego wyznaczania (indeks Youdena lub metoda stycznej). Szczegóły przedstawia ryc.3.

Rycina 3: Krzywe ROC oraz proponowane punkty odcięcia dla parametru rPET_{interim} (A - indeks Youdena, B - metoda stycznej).

A



B



Metody półautomatyczne oceny PET/CT

Próba połączenia idei wykorzystania zarówno pomiarów anatomicznych, jak i oceny metabolizmu glukozy w zmianach nowotworowych stały się metody półautomatyczne oceny badania PET. Zaliczamy do nich zarówno metaboliczną objętość guza (MTV), jak i całkowitą glikolizę zmian (TLG).

Nie we wszystkich dostępnych badaniach PET/CT było możliwe zastosowanie półautomatycznych metod oceny. Trzy badania wykluczono ze względu na sposób zapisu nieczytelny dla aplikacji Metavol. Spośród kolejnych, u części chorych wyznaczony punkt odcięcia był zbyt niski i do automatycznie wyznaczanych ROI (ang. Region Of Interest – obszar zainteresowania) były włączane znaczne obszary fizjologicznej dystrybucji radioznacznika – co wykluczało w praktyce wykorzystanie tej metody. W tabeli 12 podsumowano liczbę badań dostępnych do dalszej analizy (lub par badań, pozwalających na porównanie zmiany parametru).

Tabela 12: Liczba badań, dla których można było zastosować daną metodę półautomatycznej oceny.

Punkt odcięcia	SUV = 2,5	MBP	MBPmean + SD	41% SUVmax	SUV _{mean} wątroby + 3 SD
Badanie					
Wyjściowe	53	14	0	55	54
Interim	52	39	23	4	54
Zmiana TLG lub MTV	51	14	0	4	54

Do szczegółowej analizy statystycznej włączono te obserwacje, których ogólna liczba przekroczyła 50 – czyli wartość MTV i TLG przy punkcie odcięcia wynoszącym dla każdego z badań SUV = 2,5, przy punkcie odcięcia wyznaczanym odrębnie dla każdego z badań na poziomie 41% SUVmax (dla badania wyjściowego) lub na poziomie SUV_{mean} wątroby + 3 SD (zarówno dla badania wyjściowego, jak i interim) oraz – jeśli to miało

zastosowanie – różnicę wartości tych parametrów pomiędzy badaniem wyjściowym a interim. W tabeli 13 zaprezentowano zbiorczą charakterystykę ww. zmiennych.

Tabela 13: Wartości parametrów oceny półautomatycznej badania PET/CT, w zależności od przyjętych punktów odcięcia.

Parametr	Wszyscy pacjenci ^a	Brak progresji lub wznowy ^a	Progresja lub wznowa choroby ^a	Wartość p
MTV 2,5 (wyjściowe)	53 563,8 (267,3-846,6)	47 528,0 (264,1-835,3)	6 753,2 (563,8-1271,6)	0,178
TLG 2,5 (wyjściowe)	53 2587,6 (1335,1-3992,9)	47 2216,7 (1324,0-3963,3)	6 4190,5 (2587,6-5788,6)	0,169
MTV 2,5 (interim)	52 0 (0-0,79)	45 0 (0-0,58)	7 2,91 (0,33-14,90)	0,010
TLG 2,5 (interim)	52 0 (0-2,15)	45 0 (0-1,52)	7 8,14 (1,11-45,52)	0,008
ΔMTV 2,5	51 100 (99,91-100)	45 100 (99,93-100)	6 99,73 (99,33-99,98)	0,0501
ΔTLG 2,5	51 100 (99,95-100)	45 100 (99,97-100)	6 99,85 (99,15-99,99)	0,034
MTV 41% (wyjściowe)	55 92,37 (47,24-166,20)	48 82,11 (43,06-151,7)	7 152,89 (85,38-202,04)	0,110
TLG 41% (interim)	55 691,18 (389,09-1253,18)	48 621,01 (367,74-1157,15)	7 1629,26 (486,23-2141,52)	0,079
MTV wątroba (wyjściowe)	54 580,55 (226,51-829,22)	48 507,01 (208,99-817,14)	6 747,67 (613,51-1620,05)	0,103
TLG wątroba (wyjściowe)	54 2543,22 (1207,82-4073,14)	48 2307,27 (973,26-4011,86)	6 4808,21 (2709,18-7442,89)	0,038
MTV wątroba (interim)	54 0 (0-0)	47 0 (0-0)	7 6,54 (0,23-12,94)	<0,001
TLG wątroba (interim)	54 0 (0-0)	47 0 (0-0)	7 16,64 (0,733-38,13)	<0,001

Parametr	Wszyscy pacjenci ^a	Brak progresji lub wznowy ^a	Progresja lub wznowa choroby ^a	Wartość p
ΔMTV wątroba	54 100 (100-100)	48 100 (100-100)	6 99,49 (96,79-99,98)	0,004
ΔTLG wątroba	54 100 (100-100)	48 100 (100-100)	6 99,74 (97,22-99,99)	0,004

^a Pierwsza liczba w komórce oznacza liczbę poddanych analizie obserwacji, zaś kolejne wartości to mediana (rozstęp międzykwartyłowy) analizowanych zmiennych.

Statystycznie istotne różnice pomiędzy populacjami osób, które po leczeniu były wolne od progresji choroby oraz tych, które doświadczyły wznowy lub progresji osiągnięto przy progach odcięcia dla pomiarów półautomatycznych na poziomie:

- wartości bezwzględnej SUV wynoszącej 2,5 w badaniu interim (p dla MTV i TLG wyniosło odpowiednio 0,01 i 0,008),
- SUV na poziomie SUVmean wątroby + 3SD dla TLG w badaniu wyjściowym (p = 0,038) oraz dla obu parametrów w badaniu interim (p w obu przypadkach wyniosło poniżej 0,001),
- różnic w wartościach przedstawionych parametrów pomiędzy badaniem wyjściowym a interim (przy czym w przypadku ΔMTV 2,5 - liczonego poprzez przyjęcie bezwzględnego punktu odcięcia SUV wynoszącego 2,5 - otrzymano graniczną wartość p, wynoszącą 0,0501).

Dla powyższych danych zbudowano modele jednoczynnikowej regresji logistycznej. Wyniki oceny poszczególnych parametrów, w tym wartość p modelu, iloraz szans (OR) oraz wartość miary dobroci dopasowania Akaike (AIC) przedstawiono w tabeli 14. W przyjętych modelach najniższe wartości p oraz wartości parametru AIC otrzymano w przypadku parametrów bazujących na punkcie odcięcia wyliczanym w sposób względny - osobno dla każdego z badań na podstawie wzoru SUVmean wątroby +3SD. Niskie wartości parametru odzwierciedlającego dobroć dopasowania modelu otrzymano także w przypadku zmiany tego parametru pomiędzy badaniem wyjściowym a interim – jednak wartości p tych modeli były większe niż 0,05.

Tabela 14: Wyniki jednoczynnikowej regresji logistycznej dla modeli zbudowanych dla poszczególnych parametrów oceny półautomatycznej badania PET.

Zmienne	OR (95% CI)	Wartość p	AIC
MTV 2,5 (interim)	1,12 (0,97-1,29)	0,120	38,7
TLG 2,5 (interim)	1,05 (0,99-1,11)	0,068	37,7
Δ MTV 2,5	1,2 (0,8-1,84)	0,356	40
Δ TLG 2,5	1,4 (0,79-2,55)	0,242	39,7
TLG wątroba (wyjściowe) [/1000]	1,06 (1,01-1,11)	0,02	35,2
MTV wątroba (interim)	1,55 (1,12-2,14)	0,008	31,5
TLG wątroba (interim)	1,09 (1,01-1,17)	0,025	35,6
Δ MTV wątroba [*10]	1,37 (0,99-1,9)	0,057	29,1
Δ TLG wątroba [*10]	1,39 (0,89-2,18)	0,149	30,7

DYSKUSJA

Kontrolne badanie PET/CT, wykonywane wcześniej po rozpoczęciu leczenia (interim) okazało się być silnym narzędziem prognostycznym do określania rokowania w ziarnicy złośliwej. W badaniu zaprezentowanym przez Gallamini et al. (Gallamini *et al.*, 2007), pokazano, że przewiduje ono długotrwałe wyniki leczenia znacznie lepiej, niż Międzynarodowy Wskaźnik Prognostyczny (International Prognostic Score – IPS). Stało się to inspiracją dla niniejszej pracy i źródłem poszukiwania najlepszych sposobów oceny tego badania.

Zacytowane powyżej odkrycie stało się podstawą dla badań nad skutecznością terapii dopasowywanej indywidualnie do pacjentów w zależności od wyniku badania interim PET – jej deeskalacji u pacjentów, których interim PET był ujemny oraz intensyfikacji leczenia dla chorych z pozytywnym wynikiem tego badania. Przykładem takiej strategii może być trial RATHL (risk adapted therapy in Hodgkin lymphoma), obejmujący wyjściowo przeszło 1200 pacjentów z zaawansowaną ziarnicą złośliwą, którzy wyjściowo otrzymali 2 cykle ABVD, z kontrolą i dalszą decyzją o leczeniu w badaniu interim PET/CT (Johnson *et al.*, 2016). 84% z włączonych do badania pacjentów miało ujemny wynik tego badania – losowo przydzielono ich do grupy kontynuującej chemioterapię ABVD lub z pominięciem bleomycyny dla cykli 3-6 (grupa AVD). Odsetek przeżyć wolnych od choroby (85%) oraz ogólnych przeżyć (97%) był bardzo podobny w obu grupach. Pacjenci z ramienia AVD rzadziej cierpieli z powodu gorączki neutropenicznej oraz polekowych powikłań pulmonologicznych. Spośród 16% pozostałych pacjentów, po leczeniu opartym na schemacie BEACOPP, 74,4% osiągnęło ujemny wynik kontrolnego badania PET. Trzyletni odsetek przeżyć wolnych od choroby w tej grupie wyniósł 67,5% i był znacznie wyższy, niż w badaniach, w których kontynuowano ABVD po pozytywnym wyniku interim PET. W badaniu tym też jedynie 6,5% wszystkich pacjentów z dużymi zmianami (czynnik X – obecność „bulky disease”) otrzymało konsolidacyjną radioterapię, co sugeruje, że jej pominięcie może być racjonalną strategią u osób z negatywnym wynikiem interim PET (Shanbhag and Ambinder, 2018).

Zaprojektowano też szereg badań, w których – podobnie jak u chorych leczonych w Klinice Hematologii UJCM w Krakowie - leczenie I rzutu rozpoczynano od eskBEACOPP, zmniejszając jego intensywność w przypadku dobrej wczesnej odpowiedzi na leczenie. Uważano, że takie podejście pozwala na wydatne zwiększenie szans na obejście wyjściowej chemiooporności nowotworu (Gallamini and Kostakoglu, 2012). W ostatnich latach zostały opublikowane wyniki dwóch dużych badań klinicznych zrealizowanych według tej strategii – HD 18 (badanie, do którego rekrutowani byli pacjenci z Niemiec, Austrii, Szwajcarii, Holandii i Czech, kierowane przez German Hodgkin Study Group - GHSG) oraz AHL 2011 (badanie francusko-belgijskie, zrealizowane przez Lymphoma Study Association - LYSA).

Do badania HD18 włączono wyjściowo przeszło 2000 pacjentów w wieku od 18-60 lat, z zaawansowaną postacią ziarnicy złośliwej, tj. w stopniu II B – 14% pacjentów, III lub IV (Borchmann, Goergen, *et al.*, 2017; Borchmann, Haverkamp, *et al.*, 2017). Punktem odcięcia w badaniu interim PET – oznaczającym jego dodatni wynik – był wychwyty radioznacznika przekraczający poziom puli krwi narządów śródpiersia (tj. poziom 3 w skali Deauville, poziomy 1-2 były uważane za negatywne – podobnie jak w kryteriach IHP). Pacjenci z negatywnym wynikiem badania (1013 z 1964 ocenionych w interim PET, tj. 52%) byli randomizowani i docelowo otrzymywali dwa lub cztery dodatkowe cykle eskBEACOPP (przed dokonaną w 2011 roku poprawką protokołu była też grupa, która otrzymała sześć dodatkowych cykli eskBEACOPP). Dokładną analizę w tej grupie utrudniła dodatkowo konieczność redukcji dawek leków u części pacjentów poddawanych dłuższemu reżimowi chemioterapeutycznemu. Pacjenci z dodatnim wynikiem badania interim PET otrzymali wyjściowo 6 dodatkowych cykli eskBEACOPP z dodatkiem (ramię badawcze) lub nie przeciwciała anty-CD20 rituksymabu, po poprawce protokołu zredukowane do 4 dodatkowych cykli eskBEACOPP. W badaniu wykazano, że dodatek rituksymabu nie wpływa znacząco na wyniki leczenia pacjentów i ramię to zamknięto w trakcie trwania badania. W grupie pacjentów o ujemnym wyniku badania interim PET/CT otrzymano 5-letnie przeżycia wolne od choroby i ogólne przeżycia na poziomie 91% vs. 92% oraz 98% vs. 95%, odpowiednio dla grupy poddanej deeskalacji leczenia lub nie. W tym badaniu 5-letnie przeżycia wolne od choroby w całości grupy z pozytywnym wynikiem badania interim PET były bardzo wysokie – na

poziomie 89,7% dla tych, którzy otrzymali 8xeskBEACOPP i 88,1% w grupie, która była leczona 8 cyklami eskBEACOPPP z dodatkiem rituksimabu (5-letnie ogólne przeżycie w tych grupach wyniosło odpowiednio 96,4% i 93,9%). Obserwowane PFS były wyższe, od poziomów osiągniętych w innych badaniach, z podobną definicją pozytywnego wyniku badania interim PET (Gallamini *et al.*, 2007; Lang and Crump, 2020). Około 25 % całej grupy reprezentowali pacjenci ocenieni na 3 punkty w skali Deauville w badaniu interim PET – w analizie post-hoc tej grupy zaobserwowano bardzo wysoki odsetek przeżyć wolnych od choroby (PFS) – w grupie, która otrzymała 6 cykli eskBEACOPP 3-letni PFS wyniósł 95,9% i był wyższy nawet do poziomu osiągniętego u pacjentów z negatywnym wynikiem PET/CT, tj. na poziomie 1 i 2 w skali Deauville, którzy otrzymali podobne leczenie (3-letni PFS wyniósł tu 92,2% - (Kobe *et al.*, 2018). W tym trialu uwzględniono też ewentualną konsolidacyjną terapię – u pacjentów, którzy w badaniach po zakończeniu chemioterapii mieli aktywne metabolicznie zmiany większe niż 2,5 cm.

Badanie AHL 2011 to pierwsze, jak do tej pory opublikowane badanie kliniczne obejmujące dużą grupę pacjentów, w którym porównano wyniki leczenia standardowego i modyfikowanego pod wpływem badania interim PET (Casasnovas *et al.*, 2019). Porównano w nim leczenie polegające na podaniu wyjściowo 2 cykli eskBEACOPP, po którym kontynuowano leczenie do łącznie 6 cykli esc BEACOPP lub – w ramieniu eksperymentalnym – deeskalowano je u pacjentów z negatywnym wynikiem badania PET/CT (rozumianym jako punkty 1-3 w skali Deauville) i podawano 4 cykle ABVD. Łącznie do badania włączono 823 pacjentów w wieku 16-60 lat, z ziarnicą złośliwą w stopniu zaawansowania II B (i obecnością masy przekraczającej 1/3 szerokość śródpiersia) – 12% pacjentów, III i IV. Kolejna ocena PET/CT została wykonana po 4 cyklu chemioterapii – w razie nie osiągnięcia CRM (punkty 4 i 5 w 5-PS) pacjenci byli poddawani intensywniejszemu leczeniu terapii drugiego rzutu. Dopuszczono także wykorzystanie radioterapii – w zależności od decyzji lekarza prowadzącego. Podczas okresu obserwacji, którego mediana wyniosła 50,4 miesiąca, 5-letni odsetek przeżyć wolnych od choroby i ogólnych przeżyć był zbliżony w obu grupach, osiągając odpowiednio 86% i 96% (wynosząc odpowiednio dla grupy kontrolnej i badanej 86,2 % vs. 85,7% oraz 95,6% vs. 95,9%). Udowodniono tym samym, że redukcja leczenia na

bazie interim PET u pacjentów, którzy osiągnęli zadowalającą odpowiedź metaboliczną, daje co najmniej podobne wyniki leczenia, jak 6 cykli eskBEACOPP.

W przedstawionych powyżej i opublikowanych w 2019 roku wynikach leczenia Kliniki Hematologii UJCM, które było punktem wyjścia dla niniejszego opracowania, objęto 188 pacjentów z zaawansowaną ziarnicą złośliwą (Długosz-Danecka *et al.*, 2019). W grupie tej pacjenci na stopniu zaawansowania II BX stanowili 36,7 % całej badanej populacji. W całości populacji osiągnięto poziom 10-letnich przeżyć bez progresji choroby oraz przeżywalności ogólnej na poziomie 79,3% oraz 89,4% (w tym odpowiednio 87,2% i 95% w grupie o ujemnym wyniku badania interim PET). Dopuszczano konsolidację leczenia przez IFRT – miała ona miejsce u 64 pacjentów (w tym 92,8 % z grupy o wyjściowym stopniu zaawansowania II BX).

Zbiorcze zestawienie badań klinicznych, uwzględniających deeskalację leczenia w zaawansowanej postaci ziarnicy złośliwej, w schematów rozpoczynających się od 2 cykli eskBEACOPP, przedstawiono w tabeli 15.

Tabela 15: Zestawienie badań klinicznych, uwzględniających deeskalację leczenia oraz wyników leczenia w Klinice Hematologii UJCM (zainspirowane przez (Lang and Crump, 2020), zmodyfikowane).

Badanie	Liczba pacjentów poddanych ocenie	Mediana okresu obserwacji (miesiące)	Leczenie wyjściowe	Leczenie w przypadku deeskalacji	Ujemny interim PET PFS (%)	Ujemny interim PET OS (%)
RATHL	1119	41,2	2 ABVD	4 AVD	84,4	97,6 (3 letnie)
HD18	1945	66	2 eBEACOPP	2 eBEACOPP	91,4	96,3 (5 letnie)
AHL 2011	799	50,4	2 eBEACOPP	4 ABVD	88,9	97 (5 letnie)
KH UJCM	188	124,8	2 eBEACOPP	4 ABVD	87,2	95 (10-letnie)

Interpretacja wczesnej odpowiedzi na leczenie - badania interim PET/CT

Badanie interim PET/CT stało się standardem w ocenie efektu leczenia chorych na zaawansowane postaci ziarnicy złośliwej, choć nie ustają dyskusje na temat jego wartości oraz tego, na ile można na nim polegać. Niewątpliwie nie jest w stanie zastąpić ono oceny dokonywanej na zakończenie leczenia (Mesguich, Milpied and Hindie, 2019). Opublikowano niedawno wyniki analizy Cochrane, dotyczącej interim PET/CT w badaniach, w których nie modyfikowano leczenia w zależności od wyniku interim PET/CT (Aldin *et al.*, 2020). Ujęto w niej 23 badania, w które włączono 7335 chorych ze świeżo zdiagnozowaną ziarnicą złośliwą, z których wywnioskowano, że „uczestnicy badań, którzy osiągnęli ujemny wynik badania interim PET prawdopodobnie mają dużą przewagę w rozumieniu ogólnego przeżycia (OS) w porównaniu do grupy z dodatnim wynikiem tego badania (...) (dowody o umiarkowanej pewności). W wartościach bezwzględnych oznacza to, że 900 spośród 1000 chorych z ujemnym wynikiem interim PET prawdopodobnie przeżyje dłużej niż trzy lata, w porównaniu do 585 (95% CI 356 – 757) na 1000 uczestników z dodatnim wynikiem tego badania.”

Pomimo zwiększającej się ilości dowodów na przydatność badania interim PET/CT, jego interpretacja wciąż pozostaje wyzwaniem. Historycznie pierwszą metodą oceny PET/CT w chłoniakach były kryteria określone przez Międzynarodowy Projekt Harmonizacyjny (International Harmonization Project – IHP) (Cheson *et al.*, 2007), zaproponowane wyjściowo jako metoda oceny na końcu leczenia. Zgodnie z tymi wytycznymi, przy determinowaniu pozytywnych wyników badań za punkt odcięcia służyła pula krwi narządów śródpiersia. Jednak ta metoda interpretacji badania PET/CT wykonywanego wcześniej podczas leczenia ma istotne ograniczenie – liczne wyniki fałszywie dodatnie. Z tego też powodu w obecnie trwających badaniach klinicznych z reguły preferowany jest inny punkt odcięcia – wynik PET jest uważany za negatywny, jeśli SUV zmian chłoniakowych jest niższy od SUV wątroby – tj. reprezentuje 1,2 lub 3 na skali 5-cio stopniowej, używanej w kryteriach Deauville (zwaną też skalą londyńską lub Lugano). Ogólnie rzecz biorąc, badanie PET/CT wykonywane wcześniej podczas leczenia może pokierować terapią w dwóch kierunkach – pomóc zarówno w decyzji klinicznej odnośnie deeskalacji (Gallamini *et al.*, 2014; Johnson *et al.*, 2016), jak i intensyfikacji

leczenia (Dann *et al.*, 2017; Zaucha *et al.*, 2017). Zaproponowana i wykorzystywana w przytoczonych kryteriach metoda oceny jest prosta i praktyczna, jednak w niektórych przypadkach może być myląca. Przykładowo podczas leczenia podwyższone wartości SUVmax mogą być wynikiem zarówno samego procesu nowotworowego, jak i toczących się w obrębie widocznej masy guza procesów zapalnych (Kostakoglu *et al.*, 2012). Innym czynnikiem wpływającym na ocenę tego badania może być stłuszczenie wątroby, nieraz wiążące się z zastosowanym leczeniem, i mające wpływ na zaklasyfikowanie chorego w skali 5-cio stopniowej (Salomon *et al.*, 2017).

Włączenie do oceny parametrów z anatomicznej części badania PET/CT

Aby obejść ww. ograniczenia, niektórzy z autorów zaproponowali włączenie do oceny pacjentów z chłoniakami danych z części tomograficznej badania PET/CT. Przykładem takiego postępowania może być ocena gęstości i jednorodności (pattern) guzów lub wątroby (Flechsigt *et al.*, 2017; Salomon *et al.*, 2017). W niniejszym badaniu został przyjęty inny sposób podejścia, bazujący na analizie wielkości guzów w anatomicznej części badania (tj. tomografii komputerowej).

Znaczenie całkowitej masy guza (ang. tumour burden), dla określania rokowania pacjentów z ziarnicą złośliwą, było określane już przy stagingu definiowanym przez połączenie badania fizykalnego i oceny zdjęcia rentgenowskiego klatki piersiowej, zarówno dla wczesnych (Specht *et al.*, 1988), jak i zaawansowanych postaci ziarnicy złośliwej (Specht and Nissen, 1988a, 1988b). Sposób wyrażania tego wskaźnika ewoluował – uwzględniał m. in. sumę objętości wszystkich zmian widocznych w usg i tomografii (Gobbi *et al.*, 2001), samej tomografii – w tym przed erą standardowej digitalizacji jej wyników (Torricelli *et al.*, 2004), a także dokładne pomiary objętości zmian w tomografii komputerowej (rozumiane jako suma iloczynów pól powierzchni zmian na pojedynczych skanach i grubości warstwy), znormalizowane w stosunku do powierzchni ciała (rTB, opisany przez Gobbi *et al.*, 2004). Jest to jednak wskaźnik, którego obliczenie jest żmudne i czasochłonne, praktycznie niemożliwe do wykonania przy łóżku pacjenta. Dlatego też zaczęto szukać jego uproszczonych form. Jeden z takich sposobów - suma iloczynów wymiarów prostopadłych do 6 reprezentatywnych zmian –

został uwzględniony w klasyfikacji Chesona i Lugano (Cheson *et al.*, 1999, 2014) dla oceny tomografii komputerowej.

Inną propozycją uproszczenia tego wskaźnika było uwzględnienie „tumour burden” rozumianego jako liczba zajętych regionów i ewentualnie obecność dużej zmiany – „bulky mass”. Gobbi *et al.* (2012) zaproponowali wzór służący ocenie rokowania, uwzględniający te dwa czynniki oraz IPS. Zwrócono też uwagę na fakt, że to właśnie obecność dużej masy węzłowej sama w sobie jest niekorzystnym czynnikiem rokowniczym. Willet *et al.* (1988) udowodnili, że chorzy z guzem w śródpiersiu o objętości poniżej 200mL mają lepsze rokowanie niż ci, u których masa w śródpiersiu jest większa od 200mL. Obecność „bulky disease” została także ujęta jako cecha X w modyfikacji Costwolds (Schomberg *et al.*, 1984; Lister *et al.*, 1989).

Zmienia się definicja i rozumienie tego pojęcia – w standardach podaje się, że taka masa powinna przekraczać 1/3 szerokości klatki piersiowej lub 10 cm. Wymiar ten został zaproponowany m. in przez Bradley *et al.* (1999) – prostota pomiarów gwarantuje ich powtarzalność i łatwość do wprowadzenia w klinice (Kumar *et al.*, 2016; Szczeklik *et al.*, 2018). W badaniach wykazano znaczenie rokownicze też niższych wartości tej zmiennej - przykładowo w badaniu Kumar *et al.* (2016) za „bulky mass” uznano masę mierzącą na poprzecznym lub czołowym przekroju w tomografii komputerowej ponad 7 cm. W badaniu AHL 2011, którego wyniki zaprezentowano powyżej, za taką cechę uznano masę o szerokości przekraczającej 1/3 śródpiersia (Casasnovas *et al.*, 2019) – podobną definicję, uzupełnioną o wymiar przekraczający 5 cm, można znaleźć u Klimm *et al.* (2013).

Nie tylko wyjściowa masa guza ma znaczenie rokownicze. W ziarnicy złośliwej odpowiedzi na leczenie towarzyszy zmniejszenie się wielkości guza. Nie jest to jednak zależność liniowa – u chorych po zakończeniu leczenia stosunkowo często stwierdza się obecność mas rezydualnych, zawierających tkankę włóknistą lub martwiczą (Coyle, Kostakoglu and Evens, 2016). W kryteriach IWG (zwanym też kryteriami Chesona), obowiązujących przed upowszechnieniem się badania PET, ocena odpowiedzi na leczenie bazowała na zmianach wielkości guza/guzów, wyrażonych przez pomiar dwóch wymiarów na przekrojach poprzecznych do sześciu reprezentatywnych zmian w badaniu TK (Cheson *et al.*, 1999). Powrót do kryteriów anatomicznych, w postaci uproszczonej

reprezentowanych przez sumę pojedynczych wymiarów do trzech zmian nowotworowych, obserwuje się także w niedawno opublikowanych kryteriach RECIL (Younes *et al.*, 2017). Twórcy ostatnich kryteriów, wzorując się na kryteriach RECIST, podkreślają także znaczenie otwarcia się na pomiary wolumetryczne, szczególnie w postaci zautomatyzowanej.

W opublikowanych badaniach, dotyczących zaawansowanej postaci ziarnicy złośliwej u dorosłych, analizę wpływu zmniejszenia się wielkości mas nowotworowych, mierzonych w części tomograficznej badania PET/CT przeprowadzono jedynie w badaniu Kobe *et al.* (2014). Była to analiza podgrup badania HD15, przeprowadzonego przez German Hodgkin Study Group (GHSJ). W analizie dotyczącej oceny po zakończeniu leczenia, wykazano, że same dane z badania tomograficznego u chorych z pozytywnym wynikiem PET nie wystarczą do określenia dalszego rokowania. Jednakże zaobserwowano, że redukcja wielkości zmian nowotworowych o ponad 40% korespondowała z 18% zmniejszeniem ryzyka nawrotu choroby. Najlepszym wyróżnikiem złego rokowania było połączenie pozytywnego wyniku badania interim PET oraz niewielkiej redukcji wielkości guza/guzów.

Pozostałe opublikowane badania odnoszą się do innych podgrup niż populacja oceniana szczegółowo w niniejszej pracy:

- Kostakoglu *et al.* (2012) sprawdzili wpływ zmiany wartości SPD pomiędzy badaniem wyjściowym a interim na predykcję wyniku odpowiedzi na leczenie u 88 chorych na wczesne postaci ziarnicy złośliwej. Przedstawiono wyniki połączenia kryteriów IHP oraz obniżki wartości SPD na poziomie 65% - najwyższe wartości predykcyjne zostały osiągnięte w grupach, w których obie wartości wskazywały na dodatni lub ujemny wynik badania.
- Ilivitzki *et al.* (2013) zrealizowali badanie w populacji pediatrycznej, u 34 dzieci z ziarnicą złośliwą. Badacze doszli do wniosku, że informacje z części tomograficznej badania interim PET/CT (realizowanego przez niskodawkową tomografię) uzupełniają wnioski z interpretacji badania PET. Jednak w tym badaniu nie zaprezentowano żadnych bardziej szczegółowych rekomendacji.

Zastosowane w bieżącym badaniu pomiary bazują zarówno na ocenie obrazów wykonanych w przekrojach poprzecznych (aksjalnych) – miarami je reprezentującymi są suma najdłuższych wymiarów (SLD) lub suma iloczynów prostokątnych wymiarów (SPD), jak i na uproszczonych pomiarach objętości tych mas chłoniakowych, które uwzględniono w obliczeniu SPD (pomiary te bazowały na założeniu, że masy węzłowe bardzo często przyjmują kształt elipsoidy). Pomiarów dokonano w dwóch istotnych momentach – bezpośrednio przed rozpoczęciem leczenia (badanie wyjściowe, baseline), oraz wcześniej podczas leczenia (po drugim lub trzecim cyklu chemioterapii eskBEACOPP – jako badanie interim). Włącznie ze zmianami tych wymiarów w odpowiedzi na podane leczenie, otrzymano dziewięć zmiennych, które wzięto pod uwagę.

Choć kryteria oceny odpowiedzi na leczenie bazujące na samym badaniu tomograficznym zostały uznane za niewystarczające (Kostakoglu *et al.*, 2012), wszystkie zbadane zmienne różniły się znacząco pomiędzy grupą pacjentów z korzystnym wynikiem leczenia oraz progresją lub wznową choroby. Potwierdziły to wyniki jednoczynnikowej regresji logistycznej. Wystąpiła silna korelacja pomiędzy wszystkimi brany pod uwagę predyktorami, co oznacza, że generalnie w branych pod uwagę badaniach aktywność metaboliczna guzów mierzona w badaniu PET jest najczęściej skojarzona ze zmianą wielkości zmian chłoniakowych. Doprowadziło to do sformułowania hipotezy, że pomiary pochodzące z części tomograficznej badania mogą być komplementarne dla oceny w skali 5-cio stopniowej Deauville – uważanej dziś za złoty standard. Wszystkie badane parametry (zarówno suma objętości, SLD i SPD w badaniu wyjściowym i interim, jak i zmiany ww. parametrów pomiędzy tymi badaniami) różniły się pomiędzy grupami pacjentów o korzystnym wyniku leczenia I rzutu i progresją lub wznową choroby w 5 lat po jego zakończeniu. Wyniki jednoczynnikowej regresji logistycznej wskazały na szczególnie wysoką wartość prognostyczną czterech pomiarów – były to ocena w skali 5-cio stopniowej Deauville oraz pomiary anatomiczne w badaniu interim – bazujące zarówno na wolumetrii mierzalnych zmian chłoniakowych, jak i sumie iloczynów ich dwóch wymiarów (SPD) lub chociażby sumie pojedynczych pomiarów (SLD).

W badanej próbie, ujemna wartość predykcyjna (NPV) w badaniu interim po zastosowaniu kryteriów oceny bazujących na PET – zarówno IHP (punktem odcięcia jest

MBP) lub kryteriów Lugano (punkt odcięcia na poziomie wychwyty wątroby, Cheson *et al.*, 2014) – była bardzo wysoka i wyniosła odpowiednio 100% lub 97,7% (tabela 9). Podobna sytuacja została zaobserwowana w przypadku czułości – z wynikami na poziomie 100% w przypadku IHP i 85,5% przy zastosowaniu kryteriów Lugano. Otrzymane wyniki są z godne z danymi literaturowymi – w opublikowanych badaniach wykorzystujących te parametry czułość badania PET wynosiła od 84,9% do 98,6% (Adams, Nievelstein and Kwee, 2015). Z drugiej strony dodatnia wartość predykcyjna była niska – 20% przy wykorzystaniu kryteriów IHP lub 50% przy Lugano. Niskie PPV zostały także zaprezentowane w przypadku poprzednich badań dedykowanych interim PET/CT (Elstrom, 2012; Adams, Nievelstein and Kwee, 2015; Cheson and Kostakoglu, 2017). Oznacza to, że choć większość pacjentów z ujemnym wynikiem interim PET osiągnie długoterminowe przeżycie wolne od choroby, to dodatnie wyniki tego badania nie oznaczają jednoznacznie niepowodzenia leczenia I rzutu. Podobny trend widoczny jest w przypadku pomiarów bazujących na części tomograficznej badania interim – takich jak SLD, SPD lub objętości. Sprawdzono zatem, czy kombinacja tych zmiennych z kryteriami PET może wpłynąć na poprawę wartości prognostycznej badania. W niniejszym badaniu sytuacja, kiedy wyniki badania były uznane za pozytywne, gdy SUVmax zmian chłoniakowych był wyższy od SUV wątroby oraz SLD było wysokie, dała najwyższą wartość PPV, przy zachowaniu bardzo dobrej NPV. Jednak czułość badania była wyższa, gdy kryteriom Lugano towarzyszyły pomiary objętości zmian. W tej sytuacji czułość sięgnęła 87,5%, przy zachowaniu bardzo dobrej swoistości (94%). Szczegóły przedstawia tabela 9 ze strony 49.

Otrzymane wyniki korespondują z danymi do tej pory opublikowanymi w literaturze. Wpływ redukcji wielkości guza/guzów w badaniu interim PET/CT u chorych na ziarnicę złośliwą, znajdujących się w grupie wysokiego ryzyka został sprawdzony w analizie podgrup badania HD15, poprowadzonym przez German Hodgkin Study Group (Kobe *et al.*, 2014). W tym badaniu redukcja wielkości zmian nowotworowych większa niż 40% korespondowała z 18% obniżeniem ryzyka nawrotu choroby. Najlepsze rozróżnienie pomiędzy grupą ze złym i dobrym rokowaniem dawało połączenie zarówno pozytywnego wyniku PET/CT, jak i niewielkiej redukcji wielkości guza/guzów. Podobne badanie, jednak u chorych na chłoniaka Hodgkina we wczesnych stadiach, zostało

przeprowadzone Kostakoglu et al. (2012), którzy sprawdzili wpływ wartości SPD I jej zmian na predykcję odpowiedzi na leczenie. Wywnioskowali oni, że połączona analiza interim PET oraz różnicy w SPD daje lepsze wyniki, niż każdy z tych testów w pojedynkę. Przyjęte punkty odcięcia to kryteria IHP oceny PET oraz obniżka SPD na poziomie 65%. Zgodnie z tą analizą, największe wartości predykcyjne zostały otrzymane, jeśli obydwie zmienne wskazywały na pozytywny lub negatywny wynik badania.

Parametry badania PET – standardowe kryteria oceny

Jak już wspomniano, istnieją dwie główne metody oceny badań PET/CT, oparte na kryteriach wywodzących się jedynie z nuklearnej części badania. Są to: kryteria IHP (półjakościowa ocena wizualna, w której SUV docelowej zmiany jest porównywany z aktywnością puli krwi śródpiersia) i kryteria Deauville (gdzie SUV zmiany wyższy od SUV wątroby jest uważany za wskaźnik dodatniego wyniku badania). Pierwsza metoda została początkowo zaproponowana jako metoda oceny na zakończenie terapii I rzutu. Zastosowana jako metoda oceny wczesnego PET/CT ma bardzo wysoką ujemną wartość predykcyjną (NPV), wynoszącą od 84,4% do 98,6% (Adams, Nievelstein and Kwee, 2015) - co oznacza, że zdecydowana większość pacjentów z ujemnymi wynikami w badaniu interim PET/CT osiągnęłaby długoterminowy PFS. Została więc uznana za przewodnią w niektórych badaniach klinicznych, w których planowano deeskalację leczenia w zależności od wczesnej odpowiedzi na terapię. Przykładem może być wyjściowe podejście w pracy zrealizowanej przez Borchmann i in. (2017), porównującej efekty podania 6 lub 8 cykli eskBEACOPP w porównaniu z jedynie 4 cyklami tej chemioterapii w zależności od wczesnej odpowiedzi na leczenie we wczesnym badaniu PET / CT.

Jednakże wysoki spadek SUVmax zmian chłoniakowej jest trudny do osiągnięcia na wczesnym etapie leczenia. Biorąc pod uwagę kryteria IHP, w badanej w niniejszej pracy grupie, wynik badania interim PET/CT u prawie 70% pacjentów można uznać za pozytywny, co kontrastuje z uzyskanymi wynikami leczenia i sugeruje liczne fałszywie dodatnie wyniki. Jednak w innych badaniach, poświęconych wynikom leczenia u chorych z zaawansowaną ziarnicą złośliwą leczonych wyjściowo schematem eskBEACOPP,

odsetek pacjentów z dodatnim badaniem interim PET, interpretowanym w powyższy sposób, był niższy, osiągając 25% w grupie 69 pacjentów badanych przez Kedmi *et al.*, (2016) lub 48% spośród 1945 osób przebadanych za pomocą wczesnego badania PET/CT w zacytowanym powyżej badaniu Niemieckiej Grupy Badawczej Hodgkina (Borchmann, Goergen, *et al.*, 2017). Niemniej jednak Casasnovas i in. (2019) stwierdzili, że taka metoda interpretacji jest nieodpowiednia. Autorzy Ci argumentowali, że obniża ona wartość prognostyczną badania interim PET, niedoszacowując korzystnych wyników leczenia wśród pacjentów w grupie z dodatnim wynikiem tego badania.

Znacznie wyższą pozytywną wartość predykcyjną wczesnego PET/CT można osiągnąć, gdy punkt odcięcia jest ustalony na poziomie SUV wątroby – jeżeli wynik tego badania jest uważany za pozytywny przy zaklasyfikowaniu SUV zmiany jako 4 lub 5 w skali pięciostopniowej, jak w kryteriach Lugano. Ta metoda oceny została przetestowana na większej grupie pacjentów z zaawansowaną ziarnicą złośliwą w badaniu AHL2011 (Casasnovas *et al.*, 2019), oceniającym przydatność badania interim PET/CT w kierowaniu terapią, w tym – w przypadku dobrej odpowiedzi - umożliwianiu deeskalacji leczenia po pierwszych dwóch cyklach eskBEACOPP do ABVD. Strategia przedstawiona w części eksperymentalnej tego badania była identyczna z przedstawioną w niniejszym badaniu i służyła zmniejszeniu toksyczności leczenia bez pogorszenia wyniku leczenia. Istotną zaletą niniejszego badania jest dłuższy okres obserwacji, przekraczający 10 lat w porównaniu z 4,2 roku w badaniu AHL2011.

Znaczenie wyboru metody interpretacji pokazuje fakt, że dwie zaprezentowane powyżej strategię różnią się o dużą grupę pacjentów – tych, którzy w badaniu interim PET/CT zostali ocenieni na 3 w skali Deauville. Było to około 25% pacjentów w badaniu HD 18 oraz blisko 50% całości grupy w obecnym badaniu. Zgodnie z sugestią Casasnovas *et al.* (2019), cała ta grupa może zyskać na redukcji toksyczności leczenia przy przyjęciu strategii jego deeskalacji.

Zbiorcze zestawienie wyników badań u chorych z zaawansowaną ziarnicą złośliwą, wyjściowo leczonych schematem eskBEACOPP, zostało przedstawione w tabeli 16.

Tabela 16: Opublikowane dane i wyniki bieżącego badania, przedstawiające wartość prognostyczną badania interim PET/CT, wykonanego po drugim cyklu chemioterapii przy użyciu różnych metod interpretacji u pacjentów z zaawansowanym chłoniakiem Hodgkina, leczonych początkowo schematem eskBEACOPP (tabela zainspirowana przez Adamsa i Kwee (2017), zmieniona).

Badanie	Kryteria uznania badania interim PET za dodatnie	Leczenie	Liczba pacjentów z dodatnim wynikiem interim PET	Efekty leczenia pacjentów z dodatnim interim PET		Efekty leczenia pacjentów z negatywnym interim PET	
				PFS	OS	PFS	OS
Casasnovas et al., (2015)	DS 4-5	6 eskBEACOPP vs 2 eskBEACOPP+4 ABVD	49/368 (13%)	Est 2-letnie 72.9%	NR	Est 2-letnie 92.8%	NR
Kedmi et al., (2016)	IHP	2xeskBEACOPP+4ABVD	17/69 (25%)	NR	5-letnie 79%	NR	5-letnie 98%
Borchmann et al., (2017)	IHP	8xeskBEACOPP (2008-2010) lub 6xeskBEACOPP (od 2011) vs 4xeskBEACOPP	940/1945 (48 %)	Est 5-letnie 88.3%	Est 5-letnie 95.5%	Est 5-letnie 91.4%	Est 5-letnie 96.3%
Casasnovas et al., (2019)	DS 4-5	6xeskBEACOPP vs 2xeskBEACOPP+4ABVD	51/410 (12%)	5-letnie 70.7%	5-letnie 92.4%	5-letnie 88.9%	5-letnie 96.7%
Obecne badanie	IHP	6xeskBEACOPP lub 2-4 eskBEACOPP, z następową ABVD (łącznie 6 cykli chemioterapii)	40/58 (69%)	55.6%	83.3%	100%	100%
Obecne badanie	DS 4-5		14/58 (24%)	50%	78.6%	97.7%	100%
Obecne badanie	rPET 1,3		8/58 (14%)	25%	62.5%	96%	100%

Ilościowe metody interpretacji PET

Nieodłączną cechą badania PET jest fakt, że wyrażenie jego wyniku bazuje na parametrach ilościowych – natomiast zalecane obecnie sposoby oceny opierają się w dużej mierze na graficznym przedstawieniu tych wartości (Barrington and Kluge, 2017). Jednak wizualna ocena niesie ze sobą spore ryzyko błędu. W badaniu wykonanym przez Biggi et al. (2013) porównano ocenę badania interim PET po dwóch cyklach ABVD u 260 pacjentów z ziarnicą złośliwą przez 6 specjalistów medycyny nuklearnej. Ich ocena była w pełni zgodna jedynie w 82% przypadków. Rozbieżności wynikały m. in. z trudności w rozróżnieniu fizjologicznego i patologicznego wychwytu FDG (serce, pobudzony metabolicznie tłuszcz brązowy, jelita) oraz z obecności gruczolaka ślinianki, patologicznych złamań lub bardzo niskiego wychwytu w wątrobie. Po osiągnięciu finalnego konsensusu, PPV i NPV badania interim PET wyniosły odpowiednio 0,73 i 0,94.

Brakuje też powszechnie przyjętej definicji 5-go stopnia w skali 5 Deauville. Sformułowano go ogólnie, zaznaczając, że w tym przypadku SUVmax zmiany powinien być znacznie wyższy od SUVmax wątroby. W zależności od badaczy, stopień ten jest najczęściej przyjmowany, jeśli SUVmax zmiany przekracza 2 lub 3 krotnie SUVmax wątroby (Ly *et al.*, 2019).

Biorąc pod uwagę powyższe zastrzeżenia, skonkludowano, że sposobem poprawy powtarzalności oceny badania interim PET może być analiza ilościowa, bazująca na SUVmax lub SUVmean oraz pochodnych tej wielkości (Gallamini and Zwarthoed, 2018).

Wartością wyrażającą poziom wychwytu radioznacznika w PET jest SUV (ang. standard uptake value - standaryzowana wartość wychwytu), rozumiany jako maksymalny lub średni wychwyty w określonym regionie, odniesiony do podanej aktywności oraz parametrów reprezentujących wielkość pacjenta (najczęściej jego masę). Na podstawie tej ogólnej definicji wyróżniono szereg pochodnych tej wartości, wykorzystywanych w ocenie badania PET:

- -SUVmax, definiowany jako najwyższa wartość SUV w lokalizacji reprezentującej najwyższy metabolizm guza. Miara ta jest reprezentowana przez pojedynczy

voxel, a tym samym może być silnie zależna od niejednorodności guza lub stosunku sygnału do szumu w obrazie. W konsekwencji powtarzane pomiary SUVmax mogą wykazywać zmienność sięgającą od 5 do 30% u jednego pacjenta (Lodge, Chaudhry and Wahl, 2012).

- SUVmean, reprezentowany przez średnią pomiaru SUV w obrębie VOI (ang. volume of interest – obszar zainteresowania, odpowiadający np. masie węzłowej). Parametr ten jest mniej podatny na zakłócenia obrazu, jednak jego wartość jest zależna od obranego VOI, którego metody delineacji pozostają w dużej mierze subiektywne (Boellaard *et al.*, 2004).
- SUVpeak – czyli średnia aktywność zmiany w VOI o objętości 1 cm³, reprezentującego najgorętszy fragment guza. Jest to wartość zalecana przez kryteria PERCIST, dotyczące oceny PET dla guzów litych (Wahl, Jacene, Kasamon and Martin A. Lodge, 2009) – jako mniej zależna od stosunku sygnału do szumu i nie wymagająca definiowania metod wyznaczania granic guza.

Mimo prób standaryzacji tych wskaźników, nie ma możliwości bezpośredniego porównywania wartości SUV pomiędzy pacjentami (a tym bardziej pomiędzy np. badaniami klinicznymi, realizowanymi w różnych ośrodkach). Równocześnie rzetelność zarówno oceny wizualnej, jak i ilościowej może być zakłócona przez niespójne protokoły badania PET i metody rekonstrukcji obrazów (Boellaard *et al.*, 2019). Wynika to po części z faktu, że wychwyty radioznacznika przez zmianę jest wartością podlegającą bardzo wielu zmiennym – typu budowa tkankowa pacjenta i ilość jego tkanki tłuszczowej, czas od podania radioznacznika, stężenie glukozy we krwi, parametry aparatu PET i oprogramowania (Keyes, 1995; Visser, Boerman and Oyen, 2010).

Aby obejść te ograniczenia, Barrington (2017) postulował normalizację SUV zmian nowotworowych w stosunku do narządów referencyjnych, co potencjalnie pomogłoby w obejściu części źródeł różnic w wartościach bezwzględnych SUV. Takimi narządami odniesienia dla chłoniaków są płuca i wątroba, objęte kryteriami oceny PET, jak w zrewidowanych kryteriach Cheson (IHP) lub Lugano (Cheson *et al.*, 2007, 2014; Barrington *et al.*, 2014). Wśród dostępnej literatury hematologicznej przeważa ta druga możliwość (Hasenclever *et al.*, 2014; Hasenclever, Kurch and Kluge, 2014; Annunziata *et al.*, 2016). Jej wybór jest argumentowany tym, że

wychwył wątroby jest łatwy do zmierzenia, wysoce powtarzalny i wykazujący wysoki stosunek sygnału do szumu. Taki sposób podejścia może być reprezentowany m. in. przez:

- stosunek SUV_{peak} zmiany do SUV_{mean} wątroby. Parametr ten, nazwany qPET, jako pierwszy zaproponowany został w badaniu interim PET u dzieci chorujących na ziarnicę złośliwą (Hasenclever et al., 2014).
- porównanie SUV_{max} zmiany do SUV_{max} wątroby, zaproponowane po raz pierwszy do wykorzystania u pacjentów z ziarnicą złośliwą przez Annunziata et al. (2016) i nazwane przez autorów rPET. Było ono uproszczeniem tego parametru, pozwalającym na zmniejszenie ryzyka błędów i zmienności w wyznaczaniu ROI.

W literaturze można spotkać także wykorzystanie jako narządu referencyjnego puli krwi narządów śródpiersia, określane jako SUR. W poszukiwaniach literaturowych autorka niniejszej pracy nie trafiła na publikację prezentującą jej wykorzystanie w chłoniakach. Miara ta została zaproponowana i testowana przede wszystkim przy guzach litych (raku przełyku i niedrobnokomórkowym raku płuca - Bütof et al., 2015, 2018; Shin et al., 2017).

W badanej grupie najwyższe wartości predykcyjne badania interim PET uzyskano, gdy zaproponowano poziom odcięcia znacznie bardziej przekraczający wychwył wątroby w stosunku do kryteriów Deauville - sięgający 130% SUV_{max} wątroby (rPET = 1,3). Takiej metodzie towarzyszyła dodatnia wartość predykcyjna na poziomie 71,4%, natomiast ujemna wartość predykcyjna wyniosła 90,6% (tabela 16). W badaniu zrealizowanym przez Annunziata et al., (2016) oceniono wyniki interim PET/CT 68 pacjentów z chłoniakiem Hodgkina, początkowo leczonych schematem ABVD. W tym badaniu proponowana wartość odcięcia rPET była na niższym poziomie – wynoszącym 1,14, co pozwoliło na zachowanie 95% swoistości i 53% czułości. W ocenianej w niniejszym badaniu grupie punkt odcięcia na poziomie 1,3 rPET pozwala na osiągnięcie korzystniejszych wyników – 96% swoistości i 62,5% czułości, natomiast wartość odcięcia 1,05 (zaproponowana przez jedną z metod statystycznych – ryc. 3) prowadzi do 84% swoistości i 87,5% czułości.

Inną ilościową metodę interpretacji badania interim PET - qPET – jako pierwszy w odniesieniu do chłoniaków przedstawili Hasenclever *et al.* (2014), w badaniu dotyczącym populacji dzieci z chłoniakiem Hodgkina leczonych chemioterapią OEPA. Co ciekawe, poziom odcięcia, zaproponowany w tych badaniach, również wynosił 1,3 (jak w niniejszej pracy dla podobnego wskaźnika, rPET). Niedawno badanie RAPID dowiodło, że tę metodę interpretacji badania interim PET można zastosować w przypadku chłoniaka Hodgkina u dorosłych, leczonego początkowo zgodnie ze schematem ABVD (Georgi *et al.*, 2020). Z kolei Kurch *et al.* (2019) postulowali, że w populacji pediatrycznej jedynie bardzo wysoki wychwyty (qPET na poziomie 2) jest prognostyczny.

Obliczenia statystyczne w niniejszym badaniu dowiodły, że nie wystąpiły istotne różnice pomiędzy wartością prognostyczną między rPET i qPET w wyodrębnianiu grup pacjentów, którzy doświadczyli wznowy lub progresji oraz tymi, którzy pozostali wolni od choroby w 5 lat po zakończeniu leczenia I rzutu. Przyjęcie punktu odcięcia w badaniu interim na poziomie 1,3 rPET pozwoliło na lepsze niż przy pomocy kryteriów Lugano wyznaczenie grupy o korzystnym wyniku leczenia (tabela 16). Równocześnie warto pamiętać, że obliczenie rPET jest możliwe bez żadnego dodatkowego oprogramowania, na podstawie danych standardowo ocenianych u każdego pacjenta (i najczęściej zapisywanych w wyniku badania). Z drugiej strony qPET jest wartością mniej wrażliwą na potencjalne zakłócenia wartości SUV pojedynczych wokseli (np. stosunek sygnału do szumu - Barrington and Kluge, 2017; Kluge, 2017; Laffon and Marthan, 2017).

Zaproponowane zmienne wymagają dalszej oceny. Jednym z czynników, które może wpływać na ich wartość i wymaga dokładniejszej oceny, jest fakt, że SUV wątroby w badaniu interim PET może być w nieco zmieniony w stosunku do badania wyjściowego lub wykonywanego na zakończenie leczenia z powodu wpływu samej chemioterapii (Chiaravalloti *et al.*, 2014).

Inny wskaźnik ilościowy – ΔSUVmax

Kolejny wskaźnik ilościowy – ΔSUVmax reprezentuje różnicę między SUVmax zmiany nowotworowej w badaniu wyjściowym oraz w określonym momencie w trakcie terapii, na przykład wcześniej podczas leczenia (w interim PET) i jest odnoszony do

zniszczenia guza (Barrington and Kluge, 2017). W niniejszym badaniu wartość ta wykazała wysoką różnicę między grupą z korzystnym wynikiem leczenia a pacjentami z postępem lub nawrotem choroby, jednak w analizie regresji logistycznej jednoczynnikowej jej wartość (wyrażona przez parametr dobroci dopasowania AIC) okazała się nieco gorsza niż miar reprezentujących stosunek SUV zmian nowotworowych do wątroby.

Szereg badaczy dowodziło znaczenia tej zmiennej w przypadku chłoniaków nieziarnicznych, takich jak DLBCL, proponując w poszczególnych badaniach różne punkty odcięcia (Lin *et al.*, 2007; Casasnovas *et al.*, 2011; Itti *et al.*, 2013). Jednak dostępne są jedynie pojedyncze badania dotyczące zastosowania tej zmiennej w ziarnicy złośliwej. Wyniki uzyskane w niniejszym opracowaniu, wykazujące na nieco mniejszą przydatność Δ SUVmax w porównaniu do innych zmiennych, wywodzących się głównie z badania interim, pozostają do pewnego stopnia sprzeczne z interpretacją Hussien *et al.* (2015). Badacze Ci ocenili wyniki badań obrazowych w populacji 104 pacjentów pediatrycznych z chłoniakiem Hodgkina i doszli do wniosku, że Δ SUVmax ma wyższą wartość prognostyczną niż zmienne pochodzące tylko z jednego badania - takich jak wartości oceniane w badaniu przed rozpoczęciem leczenia, typu metaboliczna objętość guza (MTV) lub glikoliza całkowita zmian (TLG), a także qSUV, mierzony w badaniu interim. Podobne zmienne zostały wzięte pod uwagę również przez Milgrom *et al.*, (2019). Sprawdzili grupę mieszaną, składającą się z 59 dorosłych pacjentów z ziarnicą złośliwą na wszystkich poziomach zaawansowania i doszli do wniosku, że ze wszystkich skontrolowanych zmiennych to właśnie Δ SUVmax i wyjściowa całkowita MTV były niezależnymi czynnikami prognostycznymi. Zaproponowane przez nich punkty odcięcia tych dwóch wartości, odpowiednio na poziomie 71% i 225 ml, pozwoliły podzielić populację na trzy grupy o różnych 4-letnich PFS, osiągających odpowiednio 92%, 49% i 20% ($p < 0,0001$). Podobny poziom odcięcia Δ SUVmax (71%) zaproponował zespół Rossi *et al.*, (2014), którzy stwierdzili, że zmienna ta jest związana z wyższym PPV i NPV niż ocena badania w skali Deauville.

Według Georgi *et al.*, (2020) oraz przeglądu literatury wykonanego przez autorkę niniejszej pracy, nie opublikowano do tej pory badań wdrażających ilościową metodę

interpretacji badania PET/CT wykonywanego wcześniej podczas leczenia I rzutu, rozpoczynającego się od schematu eskBEACOPP.

Masa guza (ang. tumour burden) oceniana przy pomocy metod półautomatycznych w badaniu PET – metaboliczna objętość guza (MTV) oraz całkowita glikoliza zmiany (TLG)

Zgodnie ze słowami Gobbi (2014), „masa guza w ziarnicy to dużo więcej niż najlepszy czynnik prognostyczny”. Obecność dużej masy guza, najczęściej zlokalizowanej w śródpiersiu (ang. bulky mass) jest uznanym niekorzystnym czynnikiem rokowniczym (Schomberg *et al.*, 1984). W licznych badaniach wykazano też znaczenie całkowitej masy lub też objętości nowotworu dla dalszego rokowania pacjenta (Gallamini and Kostakoglu, 2012). Upowszechnienie pomiarów PET umożliwiło łatwe przejście od „tumour burden” mierzonego głównie ręcznie, na bazie tomografii komputerowej do półautomatycznych pomiarów aktywnych metabolicznie objętości zmian chłoniakowych – wyrażonych przez metaboliczną objętość guza (MTV) lub jej pochodną, biorącą pod uwagę zarówno objętość, jak i wychwyty zmiany - całkowitą glikolizę zmiany (TLG), będącą iloczynem MTV i SUVmean. W dotychczasowych badaniach półautomatyczne obliczenia odnoszące się do objętości guza w badaniu PET (MTV) zostały uznane za ważny czynnik prognostyczny w badaniu wyjściowym, przed rozpoczęciem leczenia (Gobbi, 2014).

Przykładowo Kanoun *et al.* (2014) stwierdzili, że metaboliczna objętość guza (MTV) mierzona w badaniu PET wykonywanym w ramach stagingu pacjenta, to niezależny czynnik prognostyczny, o wyższej wartości niż np. duży wymiar guza (ang. tumour bulk). Bazując na grupie 59 chorych z ziarnicą złośliwą na różnych stopniach zaawansowania, zaproponowali punkt odcięcia dla tego parametru na poziomie 225 ml. Z kolei w badaniu retrospektywnym, obejmującym 310 pacjentów z zaawansowaną ziarnicą złośliwą leczonych wyjściowo przy pomocy schematu eskBEACOPP, Mettler *et al.* (2019) udowodnili, że MTV w badaniu wyjściowym jest czynnikiem prognostycznym, pozwalającym na przewidzenie odpowiedzi w badaniu interim PET. Kahle *et al.*, (2019) w grupie 87 chorych z ziarnicą udowodnili, że znaczenie rokownicze mają: IPS, „bulky disease”, wyjściowe MTV i TLG, ale nie martwica obecna w obrębie guza.

Duże znaczenie predykcyjne wyjściowego MTV podkreślają badania wykonywane u chorych z innym nowotworem układu chłonnego, typowo charakteryzującym się wysokim wychwytem FDG – rozlanego chłoniaka z dużych komórek B (DLBCL - Mikhaeel *et al.*, 2016). Do bardzo daleko posuniętych wniosków doszli Gupta i Sigh (2020): porównując wartości MTV w badaniu wyjściowym oraz standardowe kryteria oceny interim PET wykazali, że ta pierwsza zmienna miała wyższą prognostyczną wartość niż kryteria obecnie przyjęte za „złoty standard”.

Wyniki osiągnięte w niniejszym badaniu nie do końca potwierdzają powyższe doniesienia literaturowe dla populacji z zaawansowaną ziarnicą złośliwą. W badanej próbie metaboliczna objętość guza oceniana w badaniu wyjściowym nie różniła się znacząco pomiędzy grupą o pomyślnym i niepomyślnym rokowaniu (p wyższe od 0,1, niezależnie od przyjętego sposobu pomiaru). Znaczenie takie dla badania wyjściowego miało dopiero powiązanie objętości guza ze średnim wychwytem FDG w obrębie zmiany, przedstawione w postaci wielkości nazywanej całkowitą glikolizą zmiany (TLG). Różnice pomiędzy populacjami o korzystnym rokowaniu i wznowie lub progresji choroby zaobserwowano w sytuacji, gdy wyliczenie tego parametru było oparte o poziom referencyjny wątroby ($SUV_{mean} + 3SD$).

Wykazano natomiast bardzo wysokie znaczenie prognostyczne półautomatycznych pomiarów objętości guza dla badania wykonywanego wcześniej podczas leczenia. Szczególnie korzystne wyniki (wyrażone przez niskie wartości p oraz kryterium informacyjnego Akaike) otrzymano w przypadku metabolicznej objętości guza liczonego na bazie poziomu referencyjnego wątroby ($SUV_{mean} + 3SD$). Wartości te były zbliżone do miar oceny standardowego sposobu oceny odpowiedzi na leczenie w ziarnicy złośliwej, przedstawionego przy pomocy skali 5-cio stopniowej Deauville. W rzeczy samej obie zmienne mają cechy wspólne – bazują na podobnym punkcie odcięcia, jakim jest SUV wątroby - jednak półautomatyczne sposoby oceny odpowiedzi na leczenie pozwalają na szybką i dokładną ocenę – nie dość, że bazują na danych ilościowych, to w dodatku minimalizują ryzyko popełnienia błędu i wpływu czynnika ludzkiego, szczególnie istotnego przy wdrożeniu oceny wizualnej (Skusa *et al.*, 2020).

Porównanie pomiędzy wynikami osiągniętymi przez różnych badaczy bardzo utrudnia fakt, że pomiary półautomatyczne parametrów związanych z metaboliczną

objętością guza/guzów w przedstawionych w literaturze badaniach były prowadzone w bardzo zróżnicowany sposób. Wartością graniczną (ang. threshold value) mógł być stały, absolutny poziom SUV (np. wynoszący 2,5, jak w opublikowanych do tej pory badaniach odnoszących się do wczesnej postaci ziarnicy złośliwej (Song *et al.*, 2013; Akhtari *et al.*, 2018; Milgrom *et al.*, 2019), jego wartość w stosunku do wartości SUV organu referencyjnego (jak zaproponowana w kryteriach odpowiedzi na leczenie dla guzów litych PERCIST (Wahl, Jacene, Kasamon and Martin A. Lodge, 2009), lub stały poziom wychwytu w stosunku do SUV max – tu najczęściej przyjmowana w chłoniakach wartość wynosiła 41% (Kanoun *et al.*, 2014; Ben Bouallègue *et al.*, 2017; Moskowitz *et al.*, 2017). Jak zauważyli Kostakoglu i Chauvie, (2018), nie istnieje obecnie konsensus w sprawie optymalnego algorytmu segmentacji lub ilościowego wskaźnika, reprezentującego rzeczywistą metaboliczną objętość guza. Dlatego też jednym z drugorzędowych celów niniejszej pracy stała się ocena przydatności różnych sposobów półautomatycznego wyznaczania objętości guza.

Wśród danych chorych, którzy spełnili ostateczne kryteria włączenia do niniejszego opracowania, badania PET/CT trzech pacjentów nie były zapisane w standardzie czytelny dla programu Metavol. Spośród pozostałych części obserwacji odrzucono ze względu na objęcie w badaniu zbyt dużej ilości tkanek cechujących się fizjologicznie podwyższonym wychwytem FDG, a przez to znacznie utrudnionej półautomatycznej delinacji guza (wciąż możliwej przy wykorzystaniu części tomograficznej badania, jednak nie spełniającej warunków oceny zautomatyzowanej). Problematycznymi sytuacjami mogło być przykładowo ścisłe przyleganie zmiany chłoniakowej do dużego obszaru cechującego się typowo wysokim wychwytem FDG (jak mięsień sercowy lub pęcherz moczowy) lub ustalenie punktu odcięcia na poziomie, który był niższy dla wielu lokalizacji od tła. Taka sytuacja miała miejsce stosunkowo często wówczas, gdy za punkt odcięcia przyjmowano zmienne wywodzące się z poziomu wychwytu puli krwi narządów śródpiersia (MBP lub jej pochodna – $MBP_{mean} + SD$, którą przyjęto w kryteriach PERCIST, obowiązujących dla guzów litych) (Wahl, Jacene, Kasamon and Martin A Lodge, 2009).

Zastosowanie stałej, absolutnej wartości odcięcia (ang. threshold value) na ustalonym poziomie SUV, np. wynoszącym 2,5, jest bardzo proste i szybkie do

zastosowania. Ocena taka cechuje się też dużą powtarzalnością i niezależnością od osoby oceniającej badanie. Metoda ta jest jednak trudna do zastosowania w przypadku guzów niejednorodnych – niedoszacowuje objętości szczególnie takich zmian, w których liczne voxele charakteryzują się wychwytem niższym od ustalonej wartości (Voltin *et al.*, 2020). W badanej próbie metoda ta pozwoliła na osiągnięcie kryteriów zautomatyzowanej oceny odpowiednio w 53 i 52 z 55 dostępnych ocenie badań wyjściowych i interim (czyli odpowiednio 96% i 95% badań). Niemniej w dalszej analizie statystycznej otrzymane wartości w badaniu wyjściowym nie różniły się znacząco w grupach bez i z progresją lub wznową choroby. Także w badaniu interim wartości tych parametrów okazały się nie być istotne statystycznie w analizie jednoczynnikowej regresji logistycznej.

Innym sposobem określania punktu odcięcia jest ustalenie stałej miary względnej, przykładowo wynoszącej 41% SUVmax. Taki sposób pomiaru pozwolił na wykonanie zautomatyzowanego obrysowania granic nowotworu u wszystkich osób z badanej populacji w badaniu wyjściowym, których badanie było zapisane w formacie zgodnym z DICOM. Niestety nie sprawdził się zupełnie w przypadku badania interim – gdzie w większości przypadków stosunek SUV zmiany do otoczenia był niski – w takim przypadku pomiar zautomatyzowany był możliwy jedynie u 4 (7,3%) chorych. W danych literaturowych podkreślana jest prostota i szybkość takiego pomiaru, niestety wiąże się on z niską czułością i niedokładnością w obrysowywaniu szczególnie w przypadku guzów o dużej niejednorodności i wysokim SUVmax (Kostakoglu and Chauvie, 2018). Jest to zgodne z obserwacją dokonaną w niniejszym badaniu - mimo dużej łatwości zastosowania w badaniu wyjściowym, wartości tego parametru nie różniły się istotnie pomiędzy grupami bez i z wznową lub progresją choroby.

Sposobem pomiaru, który wymaga nieco więcej nakładu pracy, jest odniesienie się do poziomu wychwytu organu referencyjnego – np. wątroby lub śródpiersia. W ten sposób otrzymywana jest wartość ściśle dopasowana do danego pacjenta i badania (Voltin *et al.*, 2020). W badanej próbie punkt odniesienia analogiczny do kryteriów PERCIST – reprezentujący SUVmean wątroby liczony w sposób półautomatyczny, powiększony o trzy odchylenia standardowe tej wartości – pozwoliły na wykorzystanie zautomatyzowanych metod pomiaru u bardzo dużej grupy – przeszło 98% ocenianych

osób, i to zarówno w badaniu wyjściowym, jak i interim. Parametr ten okazał się też być najdokładniejszy – związana z nim całkowita glikoliza zmian w badaniu wyjściowym rozróżniła grupy chorych w zależności od ich odpowiedzi na leczenie I rzutu, natomiast obydwie bazujące na tym pomiarach (zarówno MTV, jak i TLG), dokonane w badaniu interim cechowały się najwyższą istotnością statystyczną w prognozowaniu wystąpienia bądź nie wznowy lub progresji choroby.

OGRANICZENIA NINIEJSZEGO BADANIA ORAZ POTENCJALNE PRZESZKODY WE WDROŻENIU PROPONOWANYCH ROZWIĄZAŃ:

Jednym z ograniczeń niniejszego badania stał się dobór próby badanej. Ze względu na zawężone kryteria doboru – chorzy na szczególny typ nowotworu w określonym stadium zaawansowania, oceniani według jednego, spójnego schematu w jednym ośrodku - mimo przyjętego pięcioletniego okresu rekrutacji i włączenia wszystkich kolejnych osób spełniających kryteria doboru - badana próba jest stosunkowo nieduża. Dodatkowo duża wyleczalność nowotworu (co stanowi oczywistą korzyść dla pacjentów) przełożyła się na fakt małej liczności wyizolowanej podgrupy ze wznową lub progresją choroby. Częściowo starano się to skompensować przez długi, przeszło 10-letni okres obserwacji chorych, który z kolei jest atutem niniejszego badania. Jednakże, jak zauważyli Kostakoglu and Chauvie (2018), metodologia oceny badania PET, szczególnie poprzez wykorzystanie pomiarów zautomatyzowanych, wciąż jest w fazie ewolucji, a wnioski z opublikowanych do tej pory danych literaturowych (często opierających się na badaniach retrospektywnych na niewielkich grupach chorych, przekładających się na niewystarczającą licznosc osób z określonymi czynnikami ryzyka lub reprezentujących dany poziom zaawansowania choroby, nieraz leczonych w różny sposób) nie zawsze są spójne.

Kolejnym ograniczeniem jest fakt, że ze względu na brak dostępu do aparatu PET/CT w Krakowie w okresie, który objęto niniejszym badaniem, kontrole odbywały się w ośrodku odległym o około 500 km od kliniki, w której byli leczeni chorzy. Przełożyło się to na fakt, że kontrolowani w badaniu PET/CT byli pacjenci w stosunkowo dobrym stanie ogólnym, zdolni do zniesienia trudów podróży. Z drugiej strony, ze względu na charakterystykę populacji, u której diagnozuje się najczęściej ziarnicę złośliwą, znaczny odsetek chorych spośród wszystkich leczonych w Klinice Hematologii UJCM w Krakowie spełniał ten warunek - były to młode osoby, w 3 i 4 dekadzie życia, bez schorzeń towarzyszących i wyjściowo w dobrym lub dość dobrym stanie ogólnym.

Ze względu na wymienione powyżej ograniczenia osiągnięte wyniki, mimo, że są bardzo obiecujące, wymagają dalszych badań – w tym potwierdzenia ich wartości na większych populacjach.

Przeszkodą dla wdrożenia niektórych z proponowanych rozwiązań, szczególnie bazujących na pomiarach półautomatycznych, może być konieczność wykorzystywania odpowiedniego oprogramowania. Jednak coraz częściej nowe przeglądarki obrazów w formacie DICOM są wyposażane w narzędzia do wolumetrii czy wykonywania pomiarów półautomatycznych (choć u wielu dostawców to oprogramowanie jest dostępne jedynie za dodatkową opłatą). Równocześnie wykorzystane w niniejszej pracy oprogramowanie Metavol jest obecnie udostępniane w wersjach na komputery PC bez żadnych ograniczeń (Hirata, 2019). Możliwe jest także skorzystanie i rozwijanie go dla własnych potrzeb w wersji open source.

Równocześnie część z zaproponowanych miar (szczególnie $SPD_{interim}$, rPET czy ΔSUV_{max}) jest prosta do wdrożenia do praktyki klinicznej – bazując na podstawowych pomiarach, i tak standardowo ocenianych w trakcie opisu badań PET/CT, nie wymaga zastosowania żadnego dodatkowego sprzętu lub oprogramowania. Natomiast warunkiem koniecznym do ich wprowadzenia jest dobra komunikacja pomiędzy kliniką a pracownią PET/CT w celu osiągnięcia wzajemnego zrozumienia i wypracowania wspólnych spójnych standardów oceny.

Kolejnym ograniczeniem wprowadzenia zaproponowanych kryteriów oceny jest czynnik ludzki – mimo, że takie wspólne kryteria mają wiele zalet, podlegają też licznym źródłom błędów. Przykładowo oceny badań obrazowych przez różnych lekarzy w różnym czasie mogą wiązać się z błędami pomiaru, których wynikiem jest nieprawidłowa kategoryzacja (Skusa *et al.*, 2020). Badając powtarzalność oceny tomografii komputerowej zgodnie z kryteriami RECIST wykazano, że zmienność pomiędzy różnymi lekarzami oceniającymi badania obrazowe (ang. interobserver variability) jest znacząco wyższa, niż zmienność w ocenie różnych badań przez jednego lekarza (ang. intraobserver variability - Erasmus *et al.*, 2003). Proponowane są różne sposoby rozwiązania tego problemu – począwszy od „przypisania” danemu pacjentowi jednego lekarza (Skusa *et al.*, 2020), po zastosowanie półautomatycznych metod pomiaru, co do których wykonano wyższość w chłoniakach (Weßling *et al.*, 2014). Autorka niniejszej pracy może

zapropionować jeszcze jedną zasadę dobrej praktyki, którą stosuje w codziennej pracy klinicznej – zapisywanie na serwerach obrazów z dokonanymi pomiarami, umożliwiające odwoływanie się podczas oceny kolejnych badań do wcześniejszych sposobów wykonywania pomiarów.

Innym ograniczeniem technicznym w niniejszym badaniu oraz ocenie aktualnie wykonywanych badań PET/CT jest fakt, że tomografia komputerowa, towarzysząca badaniu PET, w zdecydowanej większości jest wykonywana bez dożylnego podania kontrastu. Jest to podyktowane m. in. negatywnym wpływem wzmocnienia kontrastowego na korekcję atenuacji części nuklearnej tego badania. Ocena takiego badania jest dużym wyzwaniem dla radiologa, szczególnie w zakresie tkanek, które niezbyt różnią się gęstością w badaniu tomograficznym (jak np. guza śródpiersia lub wnęki bezpośrednio przylegającego do pni naczyniowych). Jednakże wspólne wykorzystanie do interpretacji obrazów pochodzących z obu części badania pozwala na uzyskanie znacznie większej ilości informacji, niż sama tomografia komputerowa (Schaefer *et al.*, 2004). Praktyką pozwalającą na osiągnięcie możliwie najbardziej rzetelnych wyników w niniejszym badaniu było dołączenie do analizy szczegółowej oceny anatomii pacjenta w badaniach z różnego okresu leczenia (co z reguły wiązało się z różnicami w wielkości i kształcie guza).

PODSUMOWANIE

Zgodnie z rezultatami poszukiwań literaturowych (stan wiedzy na czerwiec 2020), niniejsza praca jest pierwszą badającą szeroko różne sposoby oceny odpowiedzi na leczenie w badaniu PET/CT wykonywanym wcześniej podczas leczenia oraz pierwszą proponującą wdrożenie ilościowych metod interpretacji badania PET/CT, wykonywanego dla zbadania wczesnej odpowiedzi na leczenie I rzutu u chorych z ziarnicą złośliwą leczonych wyjściowo wg. schematu eskBEACOPP (Georgi *et al.*, 2020). Zbadano w niej zarówno parametry wynikające z części tomograficznej badania PET/CT, jak nuklearnej. Podejście to jest zgodne z najnowszymi kryteriami oceny odpowiedzi na leczenie w chłoniakach (RECIL), uwzględniającymi zarówno wychwyty zmiany (zmian) mierzone poprzez SUV, jak i jej wielkość, wyrażoną przez pomiary oparte na obrazach tomografii komputerowej.

W badanych zmiennych ilościowych dotyczących części nuklearnej badania PET/CT wykazano szczególną wartość stosunku SUV_{max} zmiany docelowej do SUV_{wątroby} – punkt odcięcia na poziomie 1,3 posiada szczególnie wysoką wartość predykcyjną dla określania grupy pacjentów z korzystnym wynikiem leczenia I rzutu. Jest to parametr bazujący na podstawowych zmiennych mierzonych podczas oceny badania PET/CT – i tym samym bardzo łatwy do wdrożenia w codziennej praktyce klinicznej. Podobnie łatwym do wykorzystania parametrem jest $\Delta\text{SUV}_{\text{max}}$, opisujący różnicę w poziomie wychwyty FDG pomiędzy badaniem przed i w trakcie leczenia – mogącym dostarczyć istotnych wskazówek klinicznych na temat odpowiedzi metabolicznej guza na leczenie.

Innym nowatorskim sposobem interpretacji badania interim PET/CT, zaproponowanym w niniejszej pracy, jest ocena metabolicznej objętości guza (MTV). Wielkość ta, w opublikowanych do tej pory pracach, była sprawdzana przede wszystkim dla badania wyjściowego przed leczeniem (pierwszego lub drugiego rzutu). Tymczasem oparcie jej wyliczeń na poziomie odcięcia bazującym na ocenie wychwyty FDG przez wątrobę (SUV mean + 3 SD), zalecanego także do oceny dla guzów litych (PERCIST), czyni

ten pomiar prostym, możliwym do przeprowadzenia w badaniu w trakcie leczenia, o wyższej wartości predykcyjnej od standardowej oceny w skali 5-cio stopniowej oraz cechuje się większą dokładnością w ocenie od metod półjakościowych.

Pomiary anatomiczne, bazujące na badaniu TK, posiadają wartość prognostyczną w większości nowotworów (Gallamini and Zwarthoed, 2018). Nie jest to jednak prosta zależność - anatomiczna odpowiedź na leczenie w ziarnicy złośliwej z reguły ma miejsce później, niż metaboliczna (Ganeshan *et al.*, 2016). Ponadto często nawet po zakończeniu leczenia obserwowana jest obecność masy rezydualnej, która powoduje dodatkowe utrudnienie oceny badania tomograficznego. Zatem, choć wielkość guza zmierzona w tym badaniu (zarówno reprezentowana przez pojedynczy pomiar – SLD, sumę iloczynów wymiarów prostopadłych – SPD, jak również objętość) cechuje się gorszymi rezultatami w zakresie przewidywania odpowiedzi na leczenie niż standardowa ocena, zgodna z kryteriami Deauville, to jednak otrzymane wyniki wskazują na fakt, że dane te mogą służyć jako jej uzupełnienie i dostarczać cennych informacji rokowniczych, szczególnie w przypadkach wątpliwych. W badanej próbie wykazano szczególne znaczenie połączenia kryteriów Deauville oraz zmniejszenia się wielkości zmiany w badaniu interim do odpowiednio 33,5 cm² (SPD) lub blisko 50cm³ (objętość). Ze zbadanych zmiennych, pojedyncze pomiary, wskazane w kryteriach RECIL oraz pomiary wymiarów prostopadłych, w metodologii zgodnej z kryteriami Chesona, są prostsze do zastosowania niż wolumetria i nie wymagają żadnego dodatkowego oprogramowania. W przyszłości sytuacja może ulec zmianie, wraz z implementacją do rutynowej praktyki klinicznej metod automatycznej segmentacji i liczenia objętości zmian w nowotworach hematologicznych na bazie badania TK.

Zaproponowane miary wnoszą dodatkowe dane do standardowej oceny badania interim PET/CT i pozwalają na optymalizację protokołów klinicznych bazujących na deeskalacji leczenia w zaawansowanej ziarnicy złośliwej, co prowadzi do konkluzji o zrealizowaniu zakładanego celu niniejszej pracy. Takie podejście służy głównemu obecnie celowi leczenia ziarnicy złośliwej – osiągnięciu możliwie najlepszych wyników leczenia przy minimalizacji jego skutków ubocznych. Równocześnie zaprezentowane

wyniki otwierają pole do dalszych badań – szczególnie oceny ich przydatności w najnowszych metodach leczenia, w tym z zastosowaniem nowych biofarmaceutyków.

Perspektywa kontynuacji badań

Podczas realizacji niniejszego projektu wyłonił się szereg interesujących kierunków, w których badania mogą być kontynuowane, a których nie zrealizowano ze względu na przyjęte ograniczenia niniejszej pracy. W ocenie autorki pracy najważniejsze i najciekawsze z nich to:

- ocena odpowiedzi na leczenie w innych chłoniakach, cechujących się wysokim metabolizmem i silnym wychwytem FD – glukozy. Ten kierunek po części został zrealizowany w badaniu, które stało się inspiracją dla niniejszej pracy, poświęconemu ocenie odpowiedzi na leczenie w pierwotnym chłoniaku śródpiersia z komórek B (Kocurek *et al.*, 2016),
- ocena odpowiedzi na leczenie w badaniach obrazowych i funkcjonalnych w chłoniakach przy leczeniu z zastosowaniem nowych biofarmaceutyków (w tym pogłębiona ocena wiążącej się z nimi pseudoprogresji),
- ocena odpowiedzi na leczenie chłoniaków w badaniach łączących ocenę anatomiczną, funkcjonalną i podstawowe informacje histologiczne – potencjalnie możliwości takie daje połączenie oceny w badaniu PET oraz MRI,
- ocena odpowiedzi na leczenie chłoniaków z możliwie największym zmniejszeniem dawek promieniowania jonizującego dla celów diagnostycznych, zgodnie z zasadą ALARA – ang. as low as reasonably achievable (potencjalne kierunki działań - optymalizacja protokołów niskodawkowych CT w badaniu PET/CT i wykorzystanie MR do uzupełniającej oceny anatomicznej),
- ocena odpowiedzi na leczenie chłoniaków w automatycznych metodach wolumetrycznych, biorących pod uwagę nie tylko wynik badania PET, ale także informacje anatomiczne z badania tomografii komputerowej lub rezonansu magnetycznego. Warto wspomnieć, że na zastosowanie automatycznych metod wolumetrycznych otwarta jest również klasyfikacja odpowiedzi na leczenie w guzach litych – RECIST i czerpiąca z jej wkładu klasyfikacja odnosząca się do chłoniaków - RECIL (Skusa *et al.*, 2020). Dużym utrudnieniem w zastosowaniu większości do tej pory opracowanych automatycznych tomograficznych metod wolumetrycznych jest fakt, że z reguły badanie tomograficzne towarzyszące

ocenie PET jest wykonywane bez dożylnego podania kontrastu i densywność tkanek guza jest bardzo zbliżona do innych tkanek miękkich. W niniejszej pracy oraz przedstawionych danych literaturowych rozwiązaniem tego problemu była skojarzona ocena części PET i CT oraz dokładne porównanie anatomii chorego w różnych momentach leczenia. Takie podejście otwiera pole do wykorzystania metod sztucznej inteligencji (ang. AI - Artificial Intelligence).

PIŚMIENICTWO:

- Adams, H. J. A. and Kwee, T. C. (2016) 'Interim PET-CT Scan in Advanced Hodgkin's Lymphoma', *New England Journal of Medicine*, 375(10), pp. 999–1000. doi: 10.1056/NEJMc1609333.
- Adams, H. J. A. and Kwee, T. C. (2017) 'An evidence-based review on the value of interim FDG-PET in assessing response to therapy in lymphoma', *Seminars in Oncology*. doi: 10.1053/j.seminoncol.2018.01.005.
- Adams, H. J. A., Nievelstein, R. A. J. and Kwee, T. C. (2015) 'Prognostic value of interim FDG-PET in Hodgkin lymphoma: systematic review and meta-analysis', *British Journal of Haematology*, 170(3), pp. 356–366. doi: 10.1111/bjh.13441.
- Akhtari, M. *et al.* (2018) 'Re-classifying patients with early-stage Hodgkin lymphoma based on functional radiographic markers at presentation', *Blood*. *Blood*, 131(1), pp. 84–94. doi: 10.1182/blood-2017-04-773838.
- Aldin, A. *et al.* (2020) 'Interim PET-results for prognosis in adults with Hodgkin lymphoma: a systematic review and meta-analysis of prognostic factor studies.', *The Cochrane database of systematic reviews*. *Cochrane Database Syst Rev*, 13(1), p. CD012643. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/31525824> (Accessed: 25 June 2020).
- André, M. P. E. *et al.* (2017) 'Early PET response-adapted treatment in stage I and II hodgkin lymphoma: Final results of the randomized EORTC/LYSA/FIL H10 trial', *Journal of Clinical Oncology*, 35(16). doi: 10.1200/JCO.2016.68.6394.
- Annunziata, S. *et al.* (2016) 'Interim FDG-PET/CT in Hodgkin lymphoma: the prognostic role of the ratio between target lesion and liver SUVmax (rPET)', *Annals of Nuclear Medicine*. Springer Japan, 30(8), pp. 1–5. doi: 10.1007/s12149-016-1092-9.
- Barrington, S. F. *et al.* (2014) 'Role of imaging in the staging and response assessment of lymphoma: Consensus of the international conference on malignant lymphomas imaging working group', *Journal of Clinical Oncology*, 32(27), pp. 3048–3058. doi: 10.1200/JCO.2013.53.5229.

- Barrington, S. F. and Kluge, R. (2017) 'FDG PET for therapy monitoring in Hodgkin and non-Hodgkin lymphomas', *European journal of nuclear medicine and molecular imaging*. *European Journal of Nuclear Medicine and Molecular Imaging*, 44(Suppl 1), pp. 97–110. doi: 10.1007/s00259-017-3690-8.
- Barrington, S. F. and Meignan, M. (2019) 'Time to prepare for risk adaptation in lymphoma by standardizing measurement of metabolic tumor burden', *Journal of Nuclear Medicine*, 60(8), pp. 1096–1102. doi: 10.2967/jnumed.119.227249.
- Bednaruk-Młyński, E. *et al.* (2015) 'Comparison of positron emission tomography/computed tomography with classical contrast-enhanced computed tomography in the initial staging of Hodgkin lymphoma.', *Leukemia & lymphoma*. *Leuk Lymphoma*, 56(2), pp. 377–82. doi: 10.3109/10428194.2014.919635.
- Behringer, K. *et al.* (2013) 'Gonadal Function and Fertility in Survivors After Hodgkin Lymphoma Treatment Within the German Hodgkin Study Group HD13 to HD15 Trials', *Journal of Clinical Oncology*, 31(2), pp. 231–239. doi: 10.1200/JCO.2012.44.3721.
- Biggi, A. *et al.* (2013) 'International validation study for interim PET in ABVD-treated, advanced-stage hodgkin lymphoma: interpretation criteria and concordance rate among reviewers.', *J Nucl Med. Society of Nuclear Medicine*, 54(5), pp. 683–90. doi: 10.2967/jnumed.112.110890.
- Boellaard, R. *et al.* (2004) 'Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study.', *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*. *Society of Nuclear Medicine*, 45(9), pp. 1519–27. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15347719> (Accessed: 4 July 2020).
- Boellaard, R. *et al.* (2019) 'Updating PET/CT performance standards and PET/CT interpretation criteria should go hand in hand', *EJNMMI Research*. SpringerOpen, 9(1), p. 95. doi: 10.1186/s13550-019-0565-y.
- Bonadonna, G. *et al.* (1975) 'Combination chemotherapy of Hodgkin's disease with adriamycin, bleomycin, vinblastine, and imidazole carboxamide versus MOPP.', *Cancer*, 36(1), pp. 252–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/54209>

(Accessed: 17 November 2016).

Borchmann, P., Goergen, H., *et al.* (2017) 'PET-guided treatment in patients with advanced-stage Hodgkin's lymphoma (HD18): final results of an open-label, international, randomised phase 3 trial by the German Hodgkin Study Group', *Lancet*. *Lancet*, 390(10114), pp. 2790–2802. doi: 10.1016/S0140-6736(17)32134-7.

Borchmann, P., Haverkamp, H., *et al.* (2017) 'Progression-free survival of early interim PET-positive patients with advanced stage Hodgkin's lymphoma treated with BEACOPPescalated alone or in combination with rituximab (HD18): an open-label, international, randomised phase 3 study by the German Hodgkin Study Group.', *The Lancet. Oncology*. *Lancet Oncol*, 18(4), pp. 454–463. doi: 10.1016/S1470-2045(17)30103-1.

Ben Bouallègue, F. *et al.* (2017) 'Association between textural and morphological tumor indices on baseline PET-CT and early metabolic response on interim PET-CT in bulky malignant lymphomas.', *Medical physics*. *Med Phys*, 44(9), pp. 4608–4619. doi: 10.1002/mp.12349.

Bradley, A. J. *et al.* (1999) 'Assessment and significance of mediastinal bulk in Hodgkin's disease: comparison between computed tomography and chest radiography.', *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. *J Clin Oncol*, 17(8), pp. 2493–8. doi: 10.1200/JCO.1999.17.8.2493.

Bütöf, R. *et al.* (2015) 'Prognostic value of pretherapeutic tumor-to-blood standardized uptake ratio in patients with esophageal Carcinoma', *Journal of Nuclear Medicine*. *J Nucl Med*, 56(8), pp. 1150–1156. doi: 10.2967/jnumed.115.155309.

Bütöf, R. *et al.* (2018) 'Prognostic value of SUR in patients with trimodality treatment of locally advanced esophageal carcinoma.', *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*. *J Nucl Med*. doi: 10.2967/jnumed.117.207670.

Canellos, G. P. *et al.* (1992) 'Chemotherapy of Advanced Hodgkin's Disease with MOPP, ABVD, or MOPP Alternating with ABVD', *New England Journal of Medicine*. *N Engl J Med*, 327(21), pp. 1478–1484. doi: 10.1056/NEJM199211193272102.

Carbone, P. *et al.* (1971) 'Report of the Committee on Hodgkin's Disease Staging Classification.', *Cancer Res*, 31(11), pp. 1860–61.

Casasnovas, O. *et al.* (2015) 'Randomized Phase III Study Comparing an Early PET Driven Treatment De-Escalation to a Not PET-Monitored Strategy in Patients with Advanced Stages Hodgkin Lymphoma: Interim Analysis of the AHL2011 Lysa Study', *Blood*, 126(23). Available at: <http://www.bloodjournal.org/content/126/23/5777?sso-checked=true> (Accessed: 16 January 2018).

Casasnovas, R.-O. *et al.* (2011) 'SUVmax reduction improves early prognosis value of interim positron emission tomography scans in diffuse large B-cell lymphoma.', *Blood*, 118(1), pp. 37–43. doi: 10.1182/blood-2010-12-327767.

Casasnovas, R.-O. O. *et al.* (2019) 'PET-adapted treatment for newly diagnosed advanced Hodgkin lymphoma (AHL2011): a randomised, multicentre, non-inferiority, phase 3 study', *The Lancet Oncology*, 20(2), pp. 202–215. doi: 10.1016/S1470-2045(18)30784-8.

Cheson, B. B. D. *et al.* (1999) 'Report of an International Workshop to Standardize Response Criteria for Non-Hodgkin's Lymphomas', *J Clin Oncol*, 17(4), pp. 1244–1253.

Cheson, B. D. *et al.* (2007) 'Revised response criteria for malignant lymphoma', *Journal of Clinical Oncology*, 25(5), pp. 579–586. doi: 10.1200/JCO.2006.09.2403.

Cheson, B. D. *et al.* (2014) 'Recommendations for initial evaluation, staging, and response assessment of hodgkin and non-hodgkin lymphoma: The Lugano classification', *Journal of Clinical Oncology*, 32(27), pp. 3059–3067. doi: 10.1200/JCO.2013.54.8800.

Cheson, B. D. *et al.* (2016) 'Refinement of the Lugano Classification lymphoma response criteria in the era of immunomodulatory therapy.', *Blood*. *Blood*, 128(21), pp. 2489–2496. doi: 10.1182/blood-2016-05-718528.

Cheson, B. and Kostakoglu, L. (2017) 'FDG-PET for Early Response Assessment in Lymphomas: Part 1—Hodgkin Lymphoma', *Oncology (Williston Park)*, 31(1), pp. 45–49. Available at: <http://www.cancernetwork.com/oncology-journal/fdg-pet-early-response-assessment-lymphomas-part-1> (Accessed: 11 November 2018).

Chiaravalloti, A. *et al.* (2014) 'Factors affecting inpatient liver and mediastinal blood pool ¹⁸F-FDG standardized uptake value changes during ABVD chemotherapy in Hodgkin's lymphoma.', *European journal of nuclear medicine and molecular imaging*. Eur J Nucl Med Mol Imaging, 41(6), pp. 1123–32. doi: 10.1007/s00259-014-2703-0.

Connors, J. M. *et al.* (2018) 'Brentuximab Vedotin with Chemotherapy for Stage III or IV Hodgkin's Lymphoma', *New England Journal of Medicine*, 378(4), pp. 331–344. doi: 10.1056/NEJMoa1708984.

Coyle, M., Kostakoglu, L. and Evens, a. M. (2016) 'The evolving role of response-adapted PET imaging in Hodgkin lymphoma', *Therapeutic Advances in Hematology*, 7(2), pp. 108–125. doi: 10.1177/2040620715625615.

Dann, E. J. *et al.* (2017) 'Modification of initial therapy in early and advanced Hodgkin lymphoma, based on interim PET/CT is beneficial: a prospective multicentre trial of 355 patients', *British Journal of Haematology*, 178(5), pp. 709–718. doi: 10.1111/bjh.14734.

Długosz-Danecka, M. *et al.* (2019) 'Early chemotherapy de-escalation strategy in patients with advanced-stage Hodgkin lymphoma with negative positron emission tomography scan after 2 escalated BEACOPP cycles', *Polish Archives of Internal Medicine*, 129(4), pp. 259–266. doi: 10.20452/pamw.14786.

Duggan, D. B. *et al.* (2003) 'Randomized comparison of ABVD and MOPP/ABV hybrid for the treatment of advanced Hodgkin's disease: report of an intergroup trial.', *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. J Clin Oncol, 21(4), pp. 607–14. doi: 10.1200/JCO.2003.12.086.

Eisenhauer, E. A. *et al.* (2009) 'New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1).', *European journal of cancer (Oxford, England : 1990)*. Eur J Cancer, 45(2), pp. 228–47. doi: 10.1016/j.ejca.2008.10.026.

Elstrom, R. L. (2012) 'Interim response assessment for Hodgkin lymphoma: size matters', *Leuk Lymphoma*, 8194(November), pp. 10–12. doi: 10.3109/10428194.2012.701296.

Engert, A. *et al.* (2009) 'Escalated-Dose BEACOPP in the Treatment of Patients With

Advanced-Stage Hodgkin's Lymphoma: 10 Years of Follow-Up of the GHSG HD9 Study', *Journal of Clinical Oncology*, 27(27), pp. 4548–4554. doi: 10.1200/JCO.2008.19.8820.

Engert, A. *et al.* (2010) 'Reduced treatment intensity in patients with early-stage Hodgkin's lymphoma.', *The New England journal of medicine*. N Engl J Med, 363(7), pp. 640–52. doi: 10.1056/NEJMoa1000067.

Erasmus, J. *et al.* (2003) 'Interobserver and Intraobserver Variability in Measurement of Non-Small-Cell Carcinoma Lung Lesions: Implications for Assessment of Tumor Response', *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. J Clin Oncol, 21(13). doi: 10.1200/JCO.2003.01.144.

Flechsig, P. *et al.* (2017) 'Role of CT Density in PET/CT-Based Assessment of Lymphoma', *Molecular Imaging and Biology*. doi: 10.1007/s11307-017-1155-x.

Gallamini, A. *et al.* (2006) 'The predictive value of positron emission tomography scanning performed after two courses of standard therapy on treatment outcome in advanced stage Hodgkin's disease.', *Haematologica*. Haematologica, 91(4), pp. 475–81. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16585014> (Accessed: 22 June 2020).

Gallamini, A. *et al.* (2007) 'Early interim 2-[18F]fluoro-2-deoxy-D-glucose positron emission tomography is prognostically superior to international prognostic score in advanced-stage Hodgkin's lymphoma: A report from a joint Italian-Danish study', *Journal of Clinical Oncology*. J Clin Oncol, 25(24), pp. 3746–3752. doi: 10.1200/JCO.2007.11.6525.

Gallamini, A. *et al.* (2014) 'The predictive role of interim positron emission tomography for Hodgkin lymphoma treatment outcome is confirmed using the interpretation criteria of the Deauville five-point scale', *Haematologica*, 99(6), pp. 1107–1113. doi: 10.3324/haematol.2013.103218.

Gallamini, A. and Kostakoglu, L. (2012) 'Interim FDG-PET in Hodgkin lymphoma: A compass for a safe navigation in clinical trials?', *Blood*, 120(25), pp. 4913–4920. doi: 10.1182/blood-2012-03-403790.

Gallamini, A. and Zwarthoed, C. (2018) 'Interim FDG-PET Imaging in Lymphoma',

Seminars in Nuclear Medicine. Elsevier Inc., 48(1), pp. 17–27. doi: 10.1053/j.semnuclmed.2017.09.002.

Ganeshan, B. *et al.* (2016) 'CT-based texture analysis potentially provides prognostic information complementary to interim fdg-pet for patients with hodgkin's and aggressive non-hodgkin's lymphomas', *European Radiology*. *European Radiology*, pp. 1–9. doi: 10.1007/s00330-016-4470-8.

Georgi, T. W. *et al.* (2020) 'Quantitative assessment of interim PET in Hodgkin lymphoma : An evaluation of the qPET method in adult patients in the RAPID trial', *PLoS ONE*, 15(4), p. e0231027. doi: 10.1371/journal.pone.0231027.

Gobbi, P. G. *et al.* (2001) 'Image-aided estimate of tumor burden in Hodgkin's disease: evidence of its primary prognostic importance.', *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. *J Clin Oncol*, 19(5), pp. 1388–94. doi: 10.1200/JCO.2001.19.5.1388.

Gobbi, P. G. *et al.* (2012) 'Tumor burden in Hodgkin's lymphoma can be reliably estimated from a few staging parameters.', *Oncology reports*. *Oncol Rep*, 28(3), pp. 815–20. doi: 10.3892/or.2012.1892.

Gobbi, P. G. (2014) 'Tumor burden in Hodgkin's lymphoma: much more than the best prognostic factor.', *Critical reviews in oncology/hematology*. *Crit Rev Oncol Hematol*, 90(1), pp. 17–23. doi: 10.1016/j.critrevonc.2013.11.002.

Gordon, L. I. *et al.* (2013) 'Randomized phase III trial of ABVD versus Stanford V with or without radiation therapy in locally extensive and advanced-stage Hodgkin lymphoma: an intergroup study coordinated by the Eastern Cooperative Oncology Group (E2496).', *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. *J Clin Oncol*, 31(6), pp. 684–91. doi: 10.1200/JCO.2012.43.4803.

Gupta, N. and Singh, N. (2020) 'To evaluate prognostic significance of metabolic-derived tumour volume at staging 18-fluorodeoxyglucose PET-CT scan and to compare it with standardized uptake value-based response evaluation on interim 18-fluorodeoxyglucose PET-CT scan in patients of non-Hodgkin's lymphoma (diffuse large B-cell lymphoma).', *Nuclear medicine communications*. *Nucl Med Commun*, 41(4), pp.

395–404. doi: 10.1097/MNM.0000000000001159.

Hasenclever, D. *et al.* (2014) 'qPET - A quantitative extension of the Deauville scale to assess response in interim FDG-PET scans in lymphoma', *European Journal of Nuclear Medicine and Molecular Imaging*, 41(7), pp. 1301–1308. doi: 10.1007/s00259-014-2715-9.

Hasenclever, D. and Diehl, V. (1998) 'A prognostic score for advanced Hodgkin's disease. International Prognostic Factors Project on Advanced Hodgkin's Disease.', *The New England journal of medicine*. N Engl J Med, 339(21), pp. 1506–14. doi: 10.1056/NEJM199811193392104.

Hasenclever, D., Kurch, L. and Kluge, R. (2014) 'Sources of variability in FDG PET imaging and the qPET value: reply to Laffon and Marthan', *European Journal of Nuclear Medicine and Molecular Imaging*, 41, pp. 2155–2157. doi: 10.1007/s00259-014-2880-x.

Hirata, K. *et al.* (2014) 'A semi-automated technique determining the liver standardized uptake value reference for tumor delineation in FDG PET-CT', *PLoS ONE*, 9(8). doi: 10.1371/journal.pone.0105682.

Hirata, K. (2019) *Metavol - a volume measurement tool*. Available at: <https://www.metavol.org/home> (Accessed: 19 June 2020).

Hoskin, P. J. *et al.* (2009) 'Randomized comparison of the stanford V regimen and ABVD in the treatment of advanced Hodgkin's Lymphoma: United Kingdom National Cancer Research Institute Lymphoma Group Study ISRCTN 64141244.', *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. J Clin Oncol, 27(32), pp. 5390–6. doi: 10.1200/JCO.2009.23.3239.

Hussien, A. E. M. *et al.* (2015) 'FDG-PET Response Prediction in Pediatric Hodgkin's Lymphoma: Impact of Metabolically Defined Tumor Volumes and Individualized SUV Measurements on the Positive Predictive Value.', *Cancers*. Multidisciplinary Digital Publishing Institute (MDPI), 7(1), pp. 287–304. doi: 10.3390/cancers7010287.

Hutchings, M. *et al.* (2005) 'Prognostic value of interim FDG-PET after two or three cycles of chemotherapy in Hodgkin lymphoma', *Annals of Oncology*, 16(7), pp. 1160–1168. doi: 10.1093/annonc/mdi200.

Hutchings, M. *et al.* (2006) 'FDG-PET after two cycles of chemotherapy predicts treatment failure and progression-free survival in Hodgkin lymphoma', *Blood*. *Blood*, 107(1), pp. 52–59. doi: 10.1182/blood-2005-06-2252.

Ilivitzki, A. *et al.* (2013) 'Early interim FDG PET/CT prediction of treatment response and prognosis in pediatric Hodgkin disease - Added value of low-dose CT', *Pediatric Radiology*, 43(1), pp. 86–92. doi: 10.1007/s00247-012-2517-9.

Ilyas, H. *et al.* (2018) 'Defining the optimal method for measuring baseline metabolic tumour volume in diffuse large B cell lymphoma.', *European journal of nuclear medicine and molecular imaging*. *Eur J Nucl Med Mol Imaging*, 45(7), pp. 1142–1154. doi: 10.1007/s00259-018-3953-z.

Itti, E. *et al.* (2013) 'An international confirmatory study of the prognostic value of early PET/CT in diffuse large B-cell lymphoma: comparison between Deauville criteria and Δ SUVmax', *European Journal of Nuclear Medicine and Molecular Imaging*. Springer, 40(9), pp. 1312–1320. doi: 10.1007/s00259-013-2435-6.

Johnson, P. *et al.* (2016) 'Adapted treatment guided by interim PET-CT scan in advanced Hodgkin's lymphoma', *New England Journal of Medicine*. *N Engl J Med*, 374(25), pp. 2419–29. doi: 10.1056/NEJMoa1510093.

Johnson, P. W. M. *et al.* (2005) 'Comparison of ABVD and alternating or hybrid multidrug regimens for the treatment of advanced Hodgkin's lymphoma: results of the United Kingdom Lymphoma Group LY09 Trial (ISRCTN97144519).', *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. *J Clin Oncol*, 23(36), pp. 9208–18. doi: 10.1200/JCO.2005.03.2151.

Johnson, P. W. M. (2016) 'Response-adapted frontline therapy for Hodgkin lymphoma: are we there yet?', *Hematology*, 2016(1), pp. 316–322. doi: 10.1182/asheducation-2016.1.316.

Juweid, M. E. *et al.* (2007) 'Use of positron emission tomography for response assessment of lymphoma: Consensus of the imaging subcommittee of international harmonization project in lymphoma', *Journal of Clinical Oncology*, 25(5), pp. 571–578. doi: 10.1200/JCO.2006.08.2305.

- Kahle, X. U. *et al.* (2019) 'Relationship between semiquantitative 18F-fluorodeoxyglucose positron emission tomography metrics and necrosis in classical Hodgkin lymphoma.', *Scientific reports*. Nature Publishing Group, 9(1), p. 11073. doi: 10.1038/s41598-019-47453-5.
- Kanoun, S. *et al.* (2014) 'Baseline metabolic tumour volume is an independent prognostic factor in Hodgkin lymphoma', *European Journal of Nuclear Medicine and Molecular Imaging*. Springer, 41(9), pp. 1735–1743. doi: 10.1007/s00259-014-2783-x.
- Kanoun, S., Rossi, C. and Casasnovas, O. (2018) '[18F]FDG-PET/CT in hodgkin lymphoma: Current usefulness and perspectives', *Cancers*, 10(5), pp. 1–11. doi: 10.3390/cancers10050145.
- Kedmi, M. *et al.* (2016) 'High-Risk, Advanced-Stage Hodgkin Lymphoma: The Impact of Combined Escalated BEACOPP and ABVD Treatment in Patients Who Rapidly Achieve Metabolic Complete Remission on Interim FDG-PET/CT Scan', *Acta Haematologica*. Karger Publishers, 135(3), pp. 156–161. doi: 10.1159/000441962.
- Keyes, J. W. (1995) 'SUV: Standard uptake or silly useless value?', *Journal of Nuclear Medicine*, 36(10), pp. 1836–1839.
- Klimm, B. *et al.* (2013) 'Impact of risk factors on outcomes in early-stage Hodgkin's lymphoma: An analysis of international staging definitions', *Annals of Oncology*. Elsevier Masson SAS, 24(12), pp. 3070–3076. doi: 10.1093/annonc/mdt413.
- Kluge, R. (2017) 'Reply to : Laffon and Marthan B FDG PET for therapy monitoring in Hodgkin ' s and non-Hodgkin ' s lymphomas : qPET versus rPET ^'. *European Journal of Nuclear Medicine and Molecular Imaging*, pp. 2331–2332. doi: 10.1007/s00259-017-3826-x.
- Kobe, C. *et al.* (2014) 'Assessment of tumor size reduction improves outcome prediction of positron emission tomography/computed tomography after chemotherapy in advanced-stage Hodgkin lymphoma', *Journal of Clinical Oncology*, 32(17), pp. 1776–1781. doi: 10.1200/JCO.2013.53.2507.
- Kobe, C. *et al.* (2018) 'Outcome-based interpretation of early interim PET in advanced-stage Hodgkin lymphoma.', *Blood*. Blood, 132(21), pp. 2273–2279. doi: 10.1182/blood-

2018-05-852129.

Kocurek, A. *et al.* (2016) 'Primary mediastinal B-cell lymphoma - metabolic and anatomical features in 18FDG-PET/ CT and response to therapy', *Wspolczesna Onkologia*, 20(4), pp. 297–301. doi: 10.5114/wo.2016.61849.

Kostakoglu, L. *et al.* (2012) 'Interim [(18)F]fluorodeoxyglucose positron emission tomography imaging in stage I-II non-bulky Hodgkin lymphoma: would using combined positron emission tomography and computed tomography criteria better predict response than each test alone?', *Leukemia & lymphoma*, 53(11), pp. 2143–2150. doi: 10.3109/10428194.2012.676173.

Kostakoglu, L. and Chauvie, S. (2018) 'Metabolic Tumor Volume Metrics in Lymphoma', *Seminars in Nuclear Medicine*. Elsevier Inc., 48(1), pp. 50–66. doi: 10.1053/j.semnuclmed.2017.09.005.

Kreissl, S. *et al.* (2016) 'Cancer-related fatigue in patients with and survivors of Hodgkin's lymphoma: a longitudinal study of the German Hodgkin Study Group', *The Lancet Oncology*. Elsevier, 17(10), pp. 1453–1462. doi: 10.1016/S1470-2045(16)30093-6.

Kumar, A. *et al.* (2016) 'Definition of bulky disease in early stage Hodgkin lymphoma in computed tomography era: prognostic significance of measurements in the coronal and transverse planes', *Haematologica*, 101(10110), pp. 1237–1243. doi: 10.3324/haematol.2016.141846.

Kurch, L. *et al.* (2019) 'Only strongly enhanced residual FDG uptake in early response PET (Deauville 5 or qPET ≥ 2) is prognostic in pediatric Hodgkin lymphoma : Results of the GPOH-HD2002 trial', (October 2018), pp. 1–8. doi: 10.1002/pbc.27539.

Kurtz, D. M. *et al.* (2015) 'Noninvasive monitoring of diffuse large B-cell lymphoma by immunoglobulin high-throughput sequencing.', *Blood*. Blood, 125(24), pp. 3679–87. doi: 10.1182/blood-2015-03-635169.

Laffon, E. and Marthan (2017) 'FDG PET for therapy monitoring in Hodgkin ' s and non-Hodgkin ' s lymphomas : qPET versus rPET', *Eur J Nucl Med Mol Imaging*. European Journal of Nuclear Medicine and Molecular Imaging, 14, pp. 1602–1603. doi:

10.1007/s00259-017-3719-z.

Lang, N. and Crump, M. (2020) 'PET-adapted approaches to primary therapy for advanced Hodgkin lymphoma', *Ther Adv Hematol*, 11, pp. 1–15. doi: 10.1177/2040620720914490.

Lin, C. *et al.* (2007) 'Early 18F-FDG PET for prediction of prognosis in patients with diffuse large B-cell lymphoma: SUV-based assessment versus visual analysis.', *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*. *J Nucl Med*, 48(10), pp. 1626–1632. doi: 10.2967/jnumed.107.042093.

Lister, T. A. *et al.* (1989) 'Report of a committee convened to discuss the evaluation and staging of patients with Hodgkin's disease: Cotswolds meeting.', *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. *J Clin Oncol*, 7(11), pp. 1630–6. doi: 10.1200/JCO.1989.7.11.1630.

Lodge, M. A., Chaudhry, M. A. and Wahl, R. L. (2012) 'Noise Considerations for PET Quantification Using Maximum and Peak Standardized Uptake Value', *Journal of Nuclear Medicine*. Society of Nuclear Medicine, 53(7), pp. 1041–1047. doi: 10.2967/JNUMED.111.101733.

Ly, J. *et al.* (2019) 'The use of a proposed updated EARL harmonization of 18F-FDG PET-CT in patients with lymphoma yields significant differences in Deauville score compared with current EARL recommendations', *EJNMMI Research*. SpringerOpen, 9(1), p. 65. doi: 10.1186/s13550-019-0536-3.

Matasar, M. J. and Zelenetz, A. D. (2008) 'Overview of lymphoma diagnosis and management.', *Radiologic clinics of North America*. *Radiol Clin North Am*, 46(2), pp. 175–98, vii. doi: 10.1016/j.rcl.2008.03.005.

Mauch, P. M. *et al.* (1993) 'Patterns of presentation of Hodgkin disease. Implications for etiology and pathogenesis.', *Cancer*. *Cancer*, 71(6), pp. 2062–71. doi: 10.1002/1097-0142(19930315)71:6<2062::aid-cnrcr2820710622>3.0.co;2-0.

Meignan, M. *et al.* (2009) 'Report on the First International Workshop on Interim-PET-Scan in Lymphoma.', *Leukemia & lymphoma*, 50(8), pp. 1257–1260. doi: 10.1080/10428190903040048.

Meignan, M. *et al.* (2015) 'FDG PET/CT imaging as a biomarker in lymphoma', *European Journal of Nuclear Medicine and Molecular Imaging*. Eur J Nucl Med Mol Imaging, 42(4), pp. 623–633. doi: 10.1007/s00259-014-2973-6.

Merli, F. *et al.* (2016) 'Long-term results of the HD2000 trial comparing ABVD versus BEACOPP versus COPP-EBV-CAD in untreated patients with advanced hodgkin lymphoma: A study by fondazione Italiana Linfomi', *Journal of Clinical Oncology*, 34(11). doi: 10.1200/JCO.2015.62.4817.

Mesguich, C., Milpied, N. and Hindie, E. (2019) 'PET-guided, BEACOPPescalated therapy in advanced Hodgkin lymphoma', *The lancet oncology*, 20(e189). Available at: <https://www.thelancet.com/action/showPdf?pii=S1470-2045%2819%2930162-7> (Accessed: 25 June 2020).

Mettler, J. *et al.* (2019) 'Metabolic Tumor Volume for Response Prediction in Advanced-Stage Hodgkin Lymphoma', *Journal of Nuclear Medicine*, 60(2), pp. 207–211. doi: 10.2967/jnumed.118.210047.

Mikhaeel, N. G. *et al.* (2016) 'Combination of baseline metabolic tumour volume and early response on PET/CT improves progression-free survival prediction in DLBCL.', *European journal of nuclear medicine and molecular imaging*. Eur J Nucl Med Mol Imaging, 43(7), pp. 1209–19. doi: 10.1007/s00259-016-3315-7.

Milgrom, S. A. *et al.* (2019) 'A PET Radiomics Model to Predict Refractory Mediastinal Hodgkin Lymphoma', *Scientific Reports*. Nature Publishing Group, 9(1), p. 1322. doi: 10.1038/s41598-018-37197-z.

Moskowitz, A. J. *et al.* (2017) 'Prognostic significance of baseline metabolic tumor volume in relapsed and refractory Hodgkin lymphoma', *Blood*, 130(20), pp. 2196–2203. doi: 10.1182/blood-2017-06-788877.

Mounier, N. *et al.* (2014) 'ABVD (8 cycles) versus BEACOPP (4 escalated cycles \geq 4 baseline): final results in stage III-IV low-risk Hodgkin lymphoma (IPS 0-2) of the LYSA H34 randomized trial.', *Annals of oncology*. Oxford University Press, 25(8), pp. 1622–8. doi: 10.1093/annonc/mdu189.

Paes, F. M. *et al.* (2010) 'FDG PET/CT of extranodal involvement in non-Hodgkin

lymphoma and Hodgkin disease.’, *Radiographics : a review publication of the Radiological Society of North America, Inc.* *Radiographics*, 30(1), pp. 269–91. doi: 10.1148/rg.301095088.

Radford, J. *et al.* (2015) ‘Results of a Trial of PET-Directed Therapy for Early-Stage Hodgkin’s Lymphoma’, 372, pp. 1598–607. doi: 10.1056/NEJMoa1408648.

Raemaekers, J. M. M. *et al.* (2014) ‘Omitting Radiotherapy in early positron emission tomography-negative stage I/II Hodgkin lymphoma is associated with an increased risk of early relapse: Clinical results of the preplanned interim analysis of the randomized EORTC/LYSA/FIL H10 trial’, *Journal of Clinical Oncology*, 32(12), pp. 1188–1194. doi: 10.1200/JCO.2013.51.9298.

Rossi, Cédric *et al.* (2014) ‘Interim 18F-FDG PET SUVmax reduction is superior to visual analysis in predicting outcome early in Hodgkin lymphoma patients.’, *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*. Society of Nuclear Medicine, 55(4), pp. 569–73. doi: 10.2967/jnumed.113.130609.

Rossi, C *et al.* (2014) ‘Interim 18F-FDG PET SUVmax reduction is superior to visual analysis in predicting outcome early in Hodgkin lymphoma patients’, *J Nucl Med*, 55(4), pp. 569–573. doi: 10.2967/jnumed.113.130609.

Salomon, T. *et al.* (2017) ‘Assessment of alteration in liver 18F–FDG uptake due to steatosis in lymphoma patients and its impact on the Deauville score’, *European Journal of Nuclear Medicine and Molecular Imaging*. doi: 10.1007/s00259-017-3914-y.

Schaefer, N. G. *et al.* (2004) ‘Non-Hodgkin lymphoma and Hodgkin disease: Coregistered FDG PET and CT at staging and restaging - Do we need contrast-enhanced CT?’, *Radiology*. *Radiology*, 232(3), pp. 823–829. doi: 10.1148/radiol.2323030985.

Schomberg, P. J. *et al.* (1984) ‘Prognostic significance of mediastinal mass in adult Hodgkin’s disease.’, *Cancer*. *Cancer*, 53(2), pp. 324–8. doi: 10.1002/1097-0142(19840115)53:2<324::aid-cnrcr2820530225>3.0.co;2-e.

Shanbhag, S. and Ambinder, R. F. (2018) ‘Hodgkin Lymphoma: a review and update on recent progress’, *CA Cancer J Clin*. *CA Cancer J Clin*, 68(2), pp. 116–132. doi: 10.3322/caac.21438.Hodgkin.

Shin, S. *et al.* (2017) 'Prognostic Value of Tumor-to-Blood Standardized Uptake Ratio in Patients with Resectable Non-Small-Cell Lung Cancer.', *Nuclear medicine and molecular imaging*. Nucl Med Mol Imaging, 51(3), pp. 233–239. doi: 10.1007/s13139-016-0456-3.

Siegel, R. *et al.* (2014) 'Cancer statistics, 2014', *A Cancer Journal for Clinicians*, 64, pp. 9–29.

Skoetz, N. *et al.* (2013) 'Effect of initial treatment strategy on survival of patients with advanced-stage Hodgkin's lymphoma: a systematic review and network meta-analysis', *The Lancet Oncology*. Elsevier, 14(10), pp. 943–952. doi: 10.1016/S1470-2045(13)70341-3.

Skoetz, N. *et al.* (2017) 'Comparison of first-line chemotherapy including escalated BEACOPP versus chemotherapy including ABVD for people with early unfavourable or advanced stage Hodgkin lymphoma.', *The Cochrane database of systematic reviews*. Cochrane Database Syst Rev, 5(5), p. CD007941. doi: 10.1002/14651858.CD007941.pub3.

Skusa, C. *et al.* (2020) 'Criteria-Based Imaging and Response Evaluation of Lymphoma 20 Years After Cheson: What is New? – A Review of the Current Classifications', *RöFo - Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*. Rofo, p. Online ahead of print. doi: 10.1055/a-1091-8897.

Song, M.-K. *et al.* (2013) 'Metabolic tumor volume by positron emission tomography/computed tomography as a clinical parameter to determine therapeutic modality for early stage Hodgkin's lymphoma.', *Cancer science*. Cancer Sci, 104(12), pp. 1656–61. doi: 10.1111/cas.12282.

Specht, L. *et al.* (1988) 'Tumor burden as the most important prognostic factor in early stage Hodgkin's disease. Relations to other prognostic factors and implications for choice of treatment', *Cancer*. John Wiley & Sons, Ltd, 61(8), pp. 1719–1727. doi: 10.1002/1097-0142(19880415)61:8<1719::AID-CNCR2820610834>3.0.CO;2-A.

Specht, L. and Nissen, N. I. (1988a) 'Prognostic factors in Hodgkin's disease stage III with special reference to tumour burden.', *European journal of haematology*. Eur J Haematol, 41(1), pp. 80–7. doi: 10.1111/j.1600-0609.1988.tb00873.x.

- Specht, L. and Nissen, N. I. (1988b) 'Prognostic factors in Hodgkin's disease stage IV.', *European journal of haematology*. Eur J Haematol, 41(4), pp. 359–67. doi: 10.1111/j.1600-0609.1988.tb00210.x.
- Straus, D. *et al.* (2018) 'Improving outcomes with brentixumabvedotin (BV) plus chemotherapy in patients with newly diagnosed advanced stage Hodgkin lymphoma [ASCO Abstract 7534]', *J Clin Oncol*, 36(15 suppl.).
- Straus, D. J. *et al.* (2020) 'Brentuximab vedotin with chemotherapy for stage III/IV classical Hodgkin lymphoma: 3-year update of the ECHELON-1 study.', *Blood*. Blood, 135(10), pp. 735–742. doi: 10.1182/blood.2019003127.
- Subramaniam, R. M. (2017) 'Precision Medicine and PET/Computed Tomography', *PET Clinics*, 12(1), pp. 1–5. doi: 10.1016/j.cpet.2016.08.010.
- Szczeklik, A. (1938-2012) *et al.* (2018) *Interna Szczeklika : mały podręcznik 2018/19*. Medycyna Praktyczna. Available at: <https://books.google.pl/books?id=suTRvAEACAAJ&dq=szczeklik+gajewski&hl=pl&sa=X&ved=2ahUKEwj9nY3AlpDqAhWLyKQKHaoDB7YQ6AEwAnoECAEQAAQ> (Accessed: 20 June 2020).
- Torricelli, P. *et al.* (2004) 'Hodgkin's disease: a quantitative evaluation by computed tomography of tumor burden.', *Clinical imaging*. Clin Imaging, 28(4), pp. 239–44. doi: 10.1016/S0899-7071(03)00195-5.
- Townsend, W. and Linch, D. (2012) 'Hodgkin's lymphoma in adults', *Lancet*. Elsevier Ltd, 380(9844), pp. 836–847. doi: 10.1016/S0140-6736(12)60035-X.
- Vassilakopoulos, T. P. *et al.* (2015) 'Role of FDG-PET/CT in staging and first-line treatment of Hodgkin and aggressive B-cell lymphomas', *Memo - Magazine of European Medical Oncology*, 8(2), pp. 105–114. doi: 10.1007/s12254-015-0215-7.
- Visser, E. P., Boerman, O. C. and Oyen, W. J. G. (2010) 'SUV: from silly useless value to smart uptake value.', *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*. Society of Nuclear Medicine, 51(2), pp. 173–5. doi: 10.2967/jnumed.109.068411.

Viviani, S. *et al.* (1996) 'Alternating versus hybrid MOPP and ABVD combinations in advanced Hodgkin's disease: ten-year results.', *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. J Clin Oncol, 14(5), pp. 1421–30. doi: 10.1200/JCO.1996.14.5.1421.

Viviani, S. *et al.* (2011) 'ABVD versus BEACOPP for Hodgkin's lymphoma when high-dose salvage is planned.', *The New England journal of medicine*. N Engl J Med, 365(3), pp. 203–12. doi: 10.1056/NEJMoa1100340.

Voltin, C.-A. *et al.* (2020) 'FDG-PET Imaging for Hodgkin and Diffuse Large B-Cell Lymphoma-An Updated Overview.', *Cancers*. Multidisciplinary Digital Publishing Institute (MDPI), 12(3). doi: 10.3390/cancers12030601.

Wahl, R. L., Jacene, H., Kasamon, Y. and Lodge, Martin A (2009) 'From RECIST to PERCIST : Evolving Considerations for PET Response Criteria in Solid Tumors', *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*. NIH Public Access, 50(5), pp. 122–151. doi: 10.2967/jnumed.108.057307.

Wahl, R. L., Jacene, H., Kasamon, Y. and Lodge, Martin A. (2009) 'Response Criteria in Solid Tumors', *Journal of nuclear medicine*, 50(Suppl 1), pp. 122–150. doi: 10.2967/jnumed.108.057307.From.

Weßling, J. *et al.* (2014) 'Therapy Response Evaluation of Malignant Lymphoma in a Multicenter Study: Comparison of Manual and Semiautomatic Measurements in CT', *RöFo - Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*. © Georg Thieme Verlag KG, 186(08), pp. 768–779. doi: 10.1055/s-0033-1356424.

Willett, C. G. *et al.* (1988) 'Stage IA to IIB mediastinal Hodgkin's disease: three-dimensional volumetric assessment of response to treatment.', *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. J Clin Oncol, 6(5), pp. 819–24. doi: 10.1200/JCO.1988.6.5.819.

Younes, A. *et al.* (2013) 'Brentuximab vedotin combined with ABVD or AVD for patients with newly diagnosed Hodgkin's lymphoma: a phase 1, open-label, dose-escalation study.', *The Lancet. Oncology*. Lancet Oncol, 14(13), pp. 1348–56. doi: 10.1016/S1470-

2045(13)70501-1.

Younes, A. *et al.* (2017) 'International Working Group consensus response evaluation criteria in lymphoma (RECIL 2017)', in *Annals of oncology : official journal of the European Society for Medical Oncology*. doi: 10.1093/annonc/mdx097.

Zaucha, J. M. *et al.* (2017) 'The predictive role of interim PET after the first chemotherapy cycle and sequential evaluation of response to ABVD in Hodgkin's lymphoma patients—the Polish Lymphoma Research Group (PLRG) Observational Study', *Annals of Oncology*, 28(12), pp. 3051–3057. doi: 10.1093/annonc/mdx524.

ANEKS 1: ZBIORCZE ZESTAWIENIE KLASYFIKACJI ODPOWIEDZI NA LECZENIE W CHŁONIAKACH

Kryteria odpowiedzi na leczenie Międzynarodowej Grupy Roboczej (IWG) – kryteria Cheson (Cheson <i>et al.</i>, 1999; Skusa <i>et al.</i>, 2020)					
	CR	CRu	PR	SD	PD
Ogólne	Zanik wszelkich zmian chorobowych w obrazach radiologicznych		Brak nowych zmian	Redukcja <kryteria PR, bez spełnienia warunków PD	Pojawienie się nowych zmian
Węzły chłonne	Zmniejszenie: - poprzednio >1,5cm do ≤1,5cm - poprzednio wymiar GTP z przedziału 1-1,5 cm do ≤1cm lub redukcja SPD>75%	- węzeł chłonny >1,5 cm GTP pomimo redukcji SPD>75% - separacja mas węzłowych ze zmniejszeniem SPD>75%	-redukcja SPD ≥ 50% 6 najbardziej dominujących węzłów chłonnych/mas węzłowych - niepowiększenie się innych węzłów chłonnych		- powiększenie SPD≥50% zmian docelowych
Zmiany pozawęzłowe	- śledzona prawidłowej wielkości, jednorodna - powrót do prawidłowych wymiarów wszelkich innych narządów powiększonych przed leczeniem		-bez zwiększenia wymiarów wątroby i śledziony - zmniejszenie SPD≥50% w zakresie zmian w wątrobie i śledzionie		

Kryteria Międzynarodowego Projektu Harmonizacyjnego (IHP) - zrewidowane kryteria Chesona (Cheson <i>et al.</i>, 2007; Juweid <i>et al.</i>, 2007)				
	CR	PR	SD	PD
CT	Zmniejszenie: - poprzednio >1,5 cm do ≤1,5 cm -poprzednio wymiar GTP z przedziału 1-1,5 cm do ≤1 cm ->1cm SAD do ≤1cm SAD Wątroba śledziona: - prawidłowej wielkości, jednorodne	-redukcja SPD ≥ 50% 6 najbardziej dominujących węzłów chłonnych -niepowiększenie się innych węzłów chłonnych Wątroba/śledziona: -redukcja SPD≥50% -bez zwiększenia wymiaru narządu	-brak powiększenia się istniejących zmian	Spełnione przynajmniej 1 kryterium: -nowa zmiana>1,5 cm, niezależnie od osi pomiaru -wzrost SPD≥50% -SAD>1cm: wzrost GTD≥50% -SAD≤1 cm: wzrost SAD≥50% do wymiarów 1,5cmx1,5cm lub GTP>1,5cm -powiększenie się SPD≥50% zmian w wątrobie/śledzionie
PET dla chłoniaków wykazujących wychwyty FDG w PET	-ujemne, bez względu na wielkość (dla zmian ≥ 2 cm punktem odniesienia jest poziom MBP)	-dodatnie w zakresie istniejącej zmiany (zmian)	-dodatnie w zakresie istniejących zmian, bez nowych zmian	-dodatnie

Kryteria Lugano (Barrington <i>et al.</i>, 2014; Cheson <i>et al.</i>, 2014)				
	CR	PR	SD	PD
TK: odpowiedź radiologiczna	Spełnione wszystkie kryteria: -brak nowych zmian -redukcja ≤1,5cm GTD zmian docelowych -bez zmian pozawęzłowych <u>Wątroba/śledziona</u> - prawidłowej wielkości Nowe zmiany: - brak	-redukcja SPD max. 6 mierzalnych zmian docelowych (węzłowych /pozawęzłowych) ≥ 50% -zmiany zbyt małe do rzetelnego pomiaru:0,5x0,5 cm - wielkość zmian niemierzalnych określona jako 0x0cm <u>Wątroba/śledziona</u> -zmniejszenie >50% powiększonej śledziony Nowe zmiany: -brak	- redukcja SPD max. 6 mierzalnych zmian docelowych <5 0% -kryteria PD niespełnione Nowe zmiany: -brak	Spełnienie co najmniej 1 kryterium: Nieprawidłowe węzły chłonne: - GTD > 1,5cm i - wzrost PPD ≥ 50% i - wzrost GTD lub SAD o: 0,5 cm jeśli ≤2 cm, 1 cm jeśli > 2 cm <u>Wątroba/śledziona</u> - powiększenie śledziony >50% lub powiększenie długości śledziony ≥2 cm w odniesieniu do badania wyjściowego Nowe zmiany:

				- nawrót zmian, których wymiary powróciły do wartości prawidłowych - nowe zmiany węzłowe > 1,5 cm, niezależnie od osi - nowe zmiany pozawęzłowe > 1cm niezależnie od osi lub ≤1 cm, jeśli jednoznacznie powiązane z chłoniakiem
PET/CT: odpowiedź metaboliczna	Punkty 1/2/3 zgodnie ze skalą Deauville Brak nowych zmian	Punkty 4/5 zgodnie ze skalą Deauville ze zredukowanym wychwytem 18F-FDG w stosunku do badania wyjściowego i węzły rezydualne dowolnej wielkości Brak nowych zmian	Punkty 4/5 w skali Deauville Brak nowych zmian	Punkty 4/5 zgodnie ze skalą Deauville ze zwiększonym wychwytem 18F-FDG w stosunku do badania wyjściowego Nowe, aktywne i typowe dla chłoniaków zmiany

Objaśnienia skrótów:

CR – całkowita remisja (Complete Remission)

CRu – niepotwierdzona całkowita remisja (Complete Remission unconfirmed)

PR – częściowa remisja (partial remission)

SD – stabilizacja choroby (Stable Disease)

PD – progresja choroby (Progressive Disease)

GTP – największy wymiar na przekroju poprzecznym (Great Transverse Diameter)

SPD – suma iloczynu wymiarów prostopadłych zmian (Sum of the Products of the longest Perpendicular Diameters)

SAD – wymiar w osi krótkiej (Short Axis Diameter)

MBP – pula krwi narządów śródpiersia (Mediastinal Blood Pool)

TK – tomografia komputerowa

FDG – fluoro-deoksy-glukoza

PET– pozytronowa tomografia emisyjna (Positron Emission Tomography)

PET/CT – połączenie pozytronowej tomografii emisyjnej z badaniem tomografii komputerowej (Positron Emission Tomography/Computer Tomography)

SPIS TABEL

Tabela 1: Zestawienie aktualnego stanu wiedzy oraz innowacji wnoszonych przez niniejszą pracę.

Tabela 2: Stopień wyjściowego zaawansowania choroby Ann Arbor (za Szczeklik *et al.*, 2018, p. 1673, zmieniono).

Tabela 3: Niekorzystne czynniki rokownicze w ziarnicy złośliwej – IPS.

Tabela 4: Pięciostopniowa skala oceny wychwytu FDG przez zmiany chłoniakowe – zwana skalą Deauville. Punkty na skali odnoszą się do ognisk najwyższego wychwytu w miejscach wyjściowo zajętych przez chorobę (Gallamini and Zwarthoed, 2018).

Tabela 5: Wyjściowa charakterystyka pacjentów leczonych w Klinice Hematologii SU w Krakowie wg. schematu eskBEACOPP oraz podgrupy włączonej do finalnego etapu badania.

Tabela 6: Wyjściowa charakterystyka pacjentów włączonych do badania.

Tabela 7: Wyniki analizy danych z badania PET/CT, wykonanego jako ocena wczesnej odpowiedzi na leczenie.

Tabela 8: Wyniki jednoczynnikowej regresji logistycznej oraz miara dobroci dopasowania Akaike (AIC).

Tabela 9: Podsumowanie wyników różnych metod interpretacji badania interim PET/CT – bazujących na skali 5-cio stopniowej Deauville oraz pomiarów anatomicznych, wywodzących się z części tomograficznej.

Tabela 10: Wyniki analizy parametrów badania PET.

Tabela 11: Wyniki modeli jednoczynnikowej analizy regresji logistycznej dla najbardziej obiecujących zmiennych z badania PET.

Tabela 12: Liczba badań, dla których można było zastosować daną metodę półautomatycznej oceny.

Tabela 13: Wartości parametrów oceny półautomatycznej badania PET/CT, w zależności od przyjętych punktów odcięcia.

Tabela 14: Wyniki jednoczynnikowej regresji logistycznej dla modeli zbudowanych dla poszczególnych parametrów oceny półautomatycznej badania PET.

Tabela 15: Zestawienie badań klinicznych, uwzględniających deeskalację leczenia oraz wyników leczenia w Klinice Hematologii UJCM (zainspirowane przez (Lang and Crump, 2020), zmodyfikowane).

Tabela 16: Opublikowane dane i wyniki bieżącego badania, przedstawiające wartość prognostyczną badania interim PET/CT, wykonanego po drugim cyklu chemioterapii przy użyciu różnych metod interpretacji u pacjentów z zaawansowanym chłoniakiem Hodgkina, leczonych początkowo schematem eskBEACOPP (tabela zainspirowana przez Adamsa i Kwee (2017), zmieniona).

SPIS RYCIN

Rycina 1: Sposób wyznaczania SPD – sumy iloczynów najdłuższego wymiaru zmiany na przekroju poprzecznym oraz największego wymiaru prostopadłego do niego, uwzględniająca do sześciu reprezentatywnych zmian mierzalnych.

Część tomograficzna badania PET/CT przed rozpoczęciem leczenia (baseline), pacjentka 7.

Rycina 2: Krzywe ROC dla $SLD_{interim}$ (A), $SPD_{interim}$ (B) i $V_{interim}$ (C) oraz proponowane punkty odcięcia.

Rycina 3: Krzywe ROC oraz proponowane punkty odcięcia dla parametru $rPET_{interim}$ (A - indeks Youdena, B - metoda stycznej).

STRESZCZENIE W JĘZYKU POLSKIM

Wstęp: Badanie PET/CT wykonywane wcześniej po rozpoczęciu leczenia I rzutu (ang. Interim) jest cennym narzędziem, wykorzystywanym w szeregu badań klinicznych dotyczących leczenia ziarnicy złośliwej. Celem niniejszej pracy była identyfikacja i ocena przydatności różnych sposobów oceny tego badania.

Metody: Analiza retrospektywna danych klinicznych oraz badań PET/CT: wyjściowego – przed rozpoczęciem leczenia I rzutu – i interim. Badaniem objęto 60 chorych, leczonych wyjściowo przy pomocy schematu eskBEACOPP, w razie dobrej odpowiedzi klinicznej zmienianego na ABVD. Do analizy włączono dane anatomiczne z części tomograficznej badania, takie jak SLD (suma najdłuższych wymiarów zmian), SPD (suma iloczynów prostopadłych wymiarów zmian, suma objętości, jak również zmiany z części nuklearnej badania, reprezentujące aktywność metaboliczną guza (SUVmax i SUVpeak), organów referencyjnych (wątroby i puli krwi narządów śródpiersia) oraz półautomatyczne metody pozwalające na ocenę metabolicznej objętości guza (MTV) i całkowitej glikolizy w obrębie zmian (TLG) oraz zmiany tych parametrów w czasie.

Wyniki: Podczas okresu obserwacji, którego mediana wyniosła 10,2 lat, 7 pacjentów zmarło (2 podczas leczenia I rzutu, z przyczyn innych niż progresja choroby i zostali wykluczeni z końcowej analizy) – osiągnięto poziom przeżycia całkowitego (OS) wynoszący 88,3 % i przeżycia bez progresji (PFS) 83,3 %. Najbardziej obiecującymi anatomicznymi parametrami badania PET/CT były $SPD_{interim}$ i $V_{interim}$, dla których zaproponowano punkty odcięcia odpowiednio na poziomie 33,51 cm² oraz 49,99 cm³. Połączenie tych kryteriów z 5-cio punktową skalą Deauville pozwoliło na zwiększenie dodatniej wartości predykcyjnej badania (PPV) i zachowanie jego wysokiej ujemnej wartości predykcyjnej (NPV) - wartości te wyniosły odpowiednio 86% i 96,1%, w porównaniu do 50% i 97,7% przy zastosowaniu samych kryteriów oceny PET. Dodatkowych informacji pozwalających wyróżnić osoby, u których doszło do progresji lub wznowy choroby dostarcza także stosunek SUV zmiany do SUV wątroby (wyrażony jako rPET lub qPET) oraz parametry odnoszące się do objętości metabolicznej guza (MTV i TLG) w badaniu interim, a także różnica SUV zmiany pomiędzy badaniem wyjściowym a interim.

Wnioski: Badanie PET/CT wykonywane wcześniej podczas leczenia I rzutu dostarcza cennych informacji dla prowadzenia leczenia, jednak jego interpretacja może nastęrczać trudności. Dodanie danych anatomicznych, bazujących na pomiarach wielkości zmian w badaniu interim do pomiarów SUV może służyć jako predyktor wyników leczenia, szczególnie w niejednoznacznych przypadkach.

STRESZCZENIE W JĘZYKU ANGIELSKIM

Objective: Interim PET/CT is a valuable tool used in several clinical trials regarding treatment of Hodgkin Lymphoma (HL). The objective of this work is to identify and examine the usefulness of various assessment methods of the PET/CT, performed early after the introduction of treatment (interim).

Methods: Retrospective analysis of baseline and interim PET/CT exams of 60 consecutive patients with advanced HD, treated initially with escBEACOPP, whose further therapy was either de-escalated to ABVD or continued with escBEACOPP. The analysis included anatomical data from CT, such as SLD (sum of longest diameters), SPD (sum of the products of perpendicular diameter on transverse sections), sum of lesions' volume and changes in the corresponding measures over time and data from PET – parameters referred to the metabolic tumour activity (Suvmax and SUVpeak), reference organs (mediastinal blood pool and liver and semi-automatic methods allowing for the assessment of the metabolic tumour volume (MTV) and total lesion glycolysis (TLG).

Results: During the 10,2 years follow-up, 7 patients died (2 during the initial treatment, for reasons other than progression of the lymphoma and were excluded from the further analysis) – OS reached 88,3 % and PFS 83,3 %. The most promising values from the CT were $SPD_{interim}$ and $V_{interim}$, with proposed cut-offs at the level of 33,51 cm² and 49,99 cm³, accordingly. The addition of these measures to the 5-PS in the interim exam has allowed augmentation of the PPV of the interim PET/CT exam and maintenance of its good NPV (86% and 96.1% respectively vs. 50% and 97.7% respectively, compared PET criteria alone). Additional information, allowing to distinguish people with progression or recurrence of the disease, were provided by the ratio of SUV changes to liver SUV (expressed as rPET or qPET), parameters based on the metabolic tumour volume (MTV I TLG) in the interim PET, as well as the change of SUV between baseline and interim exams.

Conclusion: Interim FDG-PET/CT is a promising modality in driving the treatment, but its interpretation is challenging. Adding anatomical data, such as measurements of the size of the lesions in the interim exam, to the SUV measures may serve as a predictor of the treatment's results, especially in ambiguous cases.