

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/395576476>

REWOLUCJA W MEDYCYNIE NUKLEARNEJ: JAK SZTUCZNA INTELIGENCJA PRZEKSZTAŁCA DIAGNOSTYKĘ I LECZENIE NOWOTWORÓW – PRZEGLĄD NAJNOWSZEJ LITERATURY

Chapter · September 2025

CITATIONS

0

READS

51

6 authors, including:



[Krzysztof Żerdziński](#)

Medical University of Silesia

15 PUBLICATIONS 22 CITATIONS

[SEE PROFILE](#)



[Julita Janiec](#)

Medical University of Silesia

11 PUBLICATIONS 4 CITATIONS

[SEE PROFILE](#)



[Kamil Jóźwik](#)

Medical University of Silesia

28 PUBLICATIONS 55 CITATIONS

[SEE PROFILE](#)



[Paweł Marek Łajczak](#)

Medical University of Silesia

123 PUBLICATIONS 107 CITATIONS

[SEE PROFILE](#)

REWOLUCJA W MEDYCYNIE NUKLEARNEJ: JAK SZTUCZNA INTELIGENCJA PRZEKSZTAŁCA DIAGNOSTYKĘ I LECZENIE NOWOTWORÓW - PRZEGLĄD NAJNOWSZEJ LITERATURY

Krzysztof Żerdziński, Michał Gałuszewski, Julita Janiec,
Kamil Józwik, Paweł Łajczak, Bartłomiej Jurek

Studenckie Koło Naukowe przy Katedrze i Zakładzie Biofizyki im. prof. Zbigniewa Religi,
Wydział Nauk Medycznych w Zabrze, Śląski Uniwersytet Medyczny w Katowicach

Abstrakt: Wstęp: Sztuczna inteligencja (AI) odgrywa kluczową rolę w medycynie nuklearnej, przyczyniając się do rewolucji w diagnostyce i leczeniu. Algorytmy uczenia maszynowego (ML) i głębokiego uczenia (DL) pozwalają na analizę danych obrazowych, takich jak PET i SPECT, z precyzją przewyższającą możliwości ludzkie. Celem pracy było przeglądowe omówienie potencjału AI w medycynie nuklearnej, w tym jej wpływu na diagnostykę, personalizację terapii i efektywność opieki zdrowotnej. Materiały i metody: Dokonano analizy 34 publikacji naukowych z ostatniego roku, wyselekcjonowanych za pomocą słów kluczowych, takich jak „AI”, „PET”, „SPECT”, „machine learning” i „deep learning”, w bazie PubMed. Ocenę przeprowadzono według schematu PICO, skupiając się na parametrach, takich jak czułość, swoistość i AUC. Zastosowano retrospektywne podejście, uwzględniając wyniki badań nad AI w różnych obszarach medycyny nuklearnej. Wyniki: AI znacząco poprawia precyzję diagnostyczną, czułość (do 97%) i swoistość (do 92%) w zastosowaniach takich jak diagnostyka raka prostaty i płuc. W obrazowaniu PET/CT modele AI osiągnęły AUC na poziomie 0,98 w przewidywaniu progresji raka prostaty oraz 0,93 w różnicowaniu zmian nowotworowych. Automatyzacja segmentacji obrazów skróciła czas analizy z godzin do minut, redukując zmienność międzyobserwacyjną. Wyniki te potwierdzają potencjał AI w personalizacji terapii oraz prognozowaniu wyników leczenia. Wnioski: AI w medycynie nuklearnej oferuje przełomowe możliwości diagnostyczne i terapeutyczne, zwiększając precyzję oraz efektywność opieki zdrowotnej. Kluczowe pozostają jednak dalsze badania nad walidacją modeli, ich interpretowalnością oraz integracją w praktyce klinicznej. Pełne wykorzystanie potencjału AI wymaga harmonizacji danych i wielośrodkowej walidacji. AI stanowi narzędzie przyszłości, które może znacząco poprawić jakość opieki medycznej.

Słowa kluczowe: sztuczna inteligencja, medycyna nuklearna, uczenie maszynowe, rak prostaty, rak płuc

Abstract: Introduction: Artificial Intelligence (AI) plays a crucial role in nuclear medicine, driving a revolution in diagnostics and treatment. Machine Learning (ML) and Deep Learning (DL) algorithms enable image data analysis, such as PET and SPECT, with accuracy surpassing human capabilities. This study aims to review the potential of AI in nuclear medicine, focusing on its impact on diagnostics, therapy personalization, and healthcare efficiency. Materials and Methods: A total of 34 scientific publications from the past year were analyzed, selected using keywords like “AI,” “PET,” “SPECT,” “machine learning,” and “deep learning” in the PubMed database. The evaluation followed the PICO framework, focusing on parameters such as sensitivity, specificity, and AUC. A retrospective approach was adopted to examine the results of AI studies across various areas of nuclear medicine. Results: AI significantly enhances diagnostic accuracy, sensitivity (up to 97%), and specificity (up to 92%) in applications such as prostate and lung cancer diagnosis. In PET/CT imaging, AI models achieved an AUC of 0.98 for predicting prostate cancer progression and 0.93 for differentiating malignant lesions. Automated image segmentation reduced analysis time from hours to minutes while decreasing interobserver variability. These findings confirm AI’s potential in therapy personalization and outcome prediction. Conclusions: AI in nuclear medicine offers groundbreaking diagnostic and therapeutic opportunities, improving precision and healthcare efficiency. However, further research is needed to validate models, ensure interpretability, and integrate AI into clinical practice. Fully harnessing AI’s potential requires data harmonization and multicenter validation. AI represents a tool of the future, capable of significantly enhancing healthcare quality.

Keywords: artificial intelligence, nuclear medicine, machine learning, prostate cancer, lung cancer

WPROWADZENIE

Sztuczna inteligencja (ang. Artificial Intelligence, AI) to technologia umożliwiająca maszynom naśladowanie inteligentnego zachowania człowieka. Termin ten został zaproponowany przez Johna McCarthy’ego w 1956 roku, ale już wcześniej Alan Turing zaprojektował test, który miał odróżniać ludzi od maszyn. AI korzysta z zaawansowanych algorytmów oraz mocy obliczeniowej, aby analizować ogromne zbiory danych w czasie rzeczywistym. W medycynie AI jest stosowana m.in. w obrazowaniu medycznym, analityce danych i wspomaganiu decyzji klinicznych. Przykładowo, AI pomaga w diagnozowaniu chorób na podstawie obrazów radiologicznych, analizie wyników badań oraz prognozowaniu wyników leczenia, co przyspiesza procesy diagnostyczne i zwiększa ich dokładność. Uczenie maszynowe (ang. Machine Learning, ML) to poddziedzina AI, wykorzystująca algorytmy do nauki z danych bez potrzeby programowania i ingerencji człowieka. Głębokie uczenie (ang. Deep Learning, DL), będące częścią ML, opiera się na sieciach neuronowych (ang. Neural Networks, NN) naśladowujących strukturę mózgu. Algorytmy DL automatycznie uczą się cech z danych, umożliwiając modelowanie złożonych, nieliniowych zależności i często przewyższając ludzkie możliwości [1-4].

Medycyna nuklearna to dziedzina medycyny, która wykorzystuje materiały promieniotwórcze i ich emisję do diagnozowania i leczenia chorób. Radiofarmaceutyki, czyli związki chemiczne znakowane radionuklidami, są podawane pacjentom drogą doustną, dożylną lub inną, aby celować w określone narządy lub procesy patologiczne. Specjalistyczne kamery rejestrują promieniowanie, co pozwala na wizualizację patologii oraz monitorowanie terapii. Najczęściej stosowane techniki diagnostyczne to pozytonowa tomografia emisyjna (ang. Positron Emission Tomography, PET) i tomografia komputerowa z użyciem pojedynczego fotonu (ang. Single-Photon Emission Computed Tomography, SPECT). W terapii radionuklidy, takie jak jod-131 czy lutet-177, wykorzystuje się do leczenia nowotworów i innych chorób. Kluczowe są również aspekty fizyczne, takie jak dobór odpowiednich osłon radiacyjnych oraz zarządzanie odpadami radioaktywnymi, co wymaga ścisłej współpracy interdyscyplinarnej [5-7].

AI w medycynie nuklearnej wykazuje ogromny potencjał, rewolucjonizując diagnostykę i leczenie poprzez automatyzację i optymalizację procesów medycznych. AI wprowadza nową jakość w obrazowaniu, np. kardiologii nuklearnej, gdzie poprawia dokładność diagnostyki, skraca czas skanowania i minimalizuje ekspozycję na promieniowanie, jednocześnie umożliwiając bardziej precyzyjne przewidywanie ryzyka chorób sercowo-naczyniowych. Wykorzystanie ML w analizie danych PET otwiera możliwości głębszej interpretacji skomplikowanych wzorców, co wspiera wcześniejsze wykrywanie chorób i personalizację terapii. Dodatkowo AI zwiększa przejrzystość systemu medycznego, umożliwiając wykrywanie oszustw w procedurach medycyny nuklearnej, co usprawnia zarządzanie zasobami i poprawia efektywność operacyjną. Te innowacje czynią AI kluczowym narzędziem przyszłości w medycynie precyzyjnej [8-10].

Celem pracy jest przegląd najnowszej literatury na temat potencjału i możliwości wykorzystania AI w medycynie nuklearnej w kontekście praktyki klinicznej.

MATERIAŁY I METODY

W celu odpowiedniego doboru literatury przygotowano kombinację słów kluczowych do wyszukiwania w bazie danych PubMed. W skład zestawu słów kluczowych weszły: Nuclear Medicine, Molecular Imaging, PET, SPECT, Positron Emission Tomography, Single Photon Emission Computed Tomography, artificial intelligence, AI, machine learning, deep learning, CNN, DL, ML, Neural Networks. Każde z tych haseł wyszukiwano z użyciem filtra [Title], aby uwzględnić w badaniu najbardziej trafne publikacje. Analizą objęto prace

opublikowane w ciągu ostatniego roku (stan na 23.12.2024), co pozwoliło dokonać przeglądu najnowszej literatury.

Uwzględniono wszystkie typy publikacji, niezależnie od języka, a w razie wątpliwości dotyczących tłumaczenia korzystano z narzędzi OpenAI. W wyniku wyszukiwania początkowo zidentyfikowano 44 publikacje, z których 40 włączono do dalszego przeglądu po ocenie tytułów. Ostatecznie do analizy wybrano 34 prace, które były zgodne z tematyką badania oraz zawierały transparentne wyniki. Pełne teksty analizowano według schematu PICO (patrz tabela 1.), koncentrując się na wartościach dotyczących dokładności, swoistości i czułości. Do przeglądu włączano wyłącznie badania bezpośrednio związane z tematem; nie prowadzono pośredniej analizy wyników innych badań.

Tabela 1. Rozwinięcie akronimu PICO, [opracowanie własne na podstawie Cochrane Polska]

P	Population	populacja włączona do badania
I	Intervention	interwencja lub leczenie zastosowane w grupie badanej
C	Comparison	interwencja lub leczenie zastosowane w grupie kontrolnej
O	Outcome	punkt końcowy/wyniki badania

Ponadto, aby prześledzić dynamikę publikacji w czasie, ustalono, że od 2002 roku w bazie PubMed opublikowano łącznie 193 prace zawierające wskazane frazy kluczowe. Analiza liczby publikacji wykazała wzrost o charakterze wykładniczym, przy czym w 2024 roku opublikowano 44 artykuły. W okresie 2014–2019 liczba publikacji zwiększyła się jedenastokrotnie. Niemniej jednak w ciągu ostatnich trzech lat zauważono spowolnienie tego trendu – w 2024 roku opublikowano mniej prac niż w 2023 roku, co może świadczyć o stopniowym nasyceniu rynku medycznego tą tematyką. W tej sytuacji przeprowadzenie podsumowania dotychczasowych osiągnięć wydaje się szczególnie uzasadnione.

WYNIKI

W niniejszej części pracy zaprezentowano kluczowe zagadnienia związane z zastosowaniem sztucznej inteligencji w medycynie nuklearnej. Na podstawie analizy najnowszych publikacji naukowych omówiono najczęściej poruszane tematy badawcze oraz najbardziej obiecujące kierunki rozwoju technologii w tej dziedzinie.

1. Precyzyjna diagnostyka i leczenie raka prostaty

Rak prostaty (ang. Prostate Cancer, PCa) to jeden z najczęstszych nowotworów u mężczyzn, w którego diagnostyce i leczeniu medycyna nuklearna odgrywa kluczową rolę dzięki technologiom, takim jak PSMA PET/CT 9 (ang. Prostate-Specific Membrane Antigen Positron Emission Tomography–Computed Tomography) i PET/MRI. Te metody umożliwiają wykrywanie zmian nowotworowych z wysoką precyzją, co wspiera personalizację terapii. AI automatyzuje analizę obrazów, redukuje zmienność między obserwatorami, standaryzuje raportowanie i wspiera prognozowanie oraz ocenę odpowiedzi na leczenie. Wykorzystanie AI w medycynie nuklearnej poprawia skuteczność diagnostyki, umożliwia wcześniejsze wykrywanie zmian i optymalizuje proces leczenia, dostosowując go do potrzeb pacjenta [11, 12].

Publikacja Tapper i wsp. [13] analizuje zastosowanie AI i radiomiki w obrazowaniu molekularnym PCa z wykorzystaniem PET/CT, w szczególności PSMA-PET. AI osiągała AUC 0,81–0,86 (Area Under Curve) w przewidywaniu Gleason Score, przerzutów i rozległości pozagruzołowej, a modele radiomiczne umożliwiały wykrywanie zmian z czułością 97% i swoistością 82% (AUC 0,98). Wykrywalność trudnych do zauważenia zmian osiągnęła AUC 0,93. AI poprawiła personalizację terapii w 50% przypadków nawrotu biochemicznego, a dokładność radiomiki w przewidywaniu progresji choroby na podstawie obrazów 18F-choline PET/CT wyniosła 67,6% (Tumor: 87%, Nodes: 82,6%, Metastasis: 72,5%). Wyniki potwierdzają potencjał AI w poprawie diagnostyki i terapii PCa, jednak wymagają walidacji na większych populacjach.

Przegląd zastosowań AI w PSMA PET/CT dla obrazowania PCa Belal i wsp. [14] wykazał, że modele AI dorównują interpretacji przez lekarzy w detekcji zmian pierwotnych i przerzutów. W detekcji węzłów chłonnych czułość AI osiągnęła 82% (77% dla lekarzy), a w ocenie obciążenia nowotworowego wykazano istotną korelację z manualnymi wynikami ($P=0,53–0,83$). Modele klasyfikacji zmian osiągały AUC 0,95–0,98, z czułością do 95% i swoistością 92%. Modele prognozujące odpowiedź na terapię [177Lu]Lu-PSMA-617 osiągnęły AUC 0,80, z czułością i swoistością 75%. Wykazano, że AI może przyspieszyć analizę obrazów, zredukować zmienność między ekspertami i dostarczać danych ilościowych, jednak potrzeba dalszej walidacji i badań nad wpływem na wyniki kliniczne, aby narzędzia AI były szerzej stosowane w praktyce klinicznej.

Przegląd systematyczny Liu i wsp. [12] oceniający zdolność AI do analizy PSMA PET w diagnostyce PCa wykazał, że AI pozwala różnicować PCa

wysokiego stopnia (International Society of Urological Pathology Grade Group, ISUP GG ≥ 3) z dokładnością 0,671–0,992, czułością 91% i swoistością 35%. W przypadku ISUP GG ≥ 4 dokładność wynosiła 0,83–0,88, czułość 89%, a swoistość 87%. AI identyfikowała zmiany niewidoczne w PSMA PET z dokładnością 87% (AUC 0,93). Modele AI wykazały potencjał w automatycznej segmentacji i pomiarze objętości guza (ang. Gross Tumor Volume, GTV), osiągając czas segmentacji ~6 s przy braku różnic w porównaniu z ręcznymi ocenami (Dice Similarity Coefficient, DSC > 0,8). Modele wykazywały zdolność do prognozowania inwazji pozaprzewodowej (AUC 0,70–0,77) i wykrywania zmian niewykrywalnych tradycyjnie. Mimo obiecujących wyników AI pozostaje narzędziem eksperymentalnym, wymagającym walidacji na większych, wieloośrodkowych populacjach przed wprowadzeniem do rutynowej praktyki klinicznej.

Badanie Huag i wsp. [15] oceniające automatyczną segmentację i detekcję zmian nowotworowych u pacjentów z PCa z nawrotem po prostatektomii wykazało, że model oparty na Konwolucyjnych Sieciach Neuronowych 3D (ang. Convolutional neural network, CNN) osiągnął DSC na poziomie 0,631 (wewnętrzny test) i 0,596 (zewnętrzny test). Czułość modelu wyniosła 72,1% wewnątrznie i 79,2% zewnątrznie, a precyzja odpowiednio 96,1% i 88,8%, co przełożyło się na F1-score na poziomie 82,4% wewnątrznie i 83,7% zewnątrznie. Model wykazał wysoką korelację z wartościami wyznaczanymi manualnie dla obciążenia guzem ($R^2 \geq 0,991$, $p < 0,05$). Automatyczna metoda umożliwia szybkie (ok. 5 minut) i powtarzalne wyznaczanie obciążenia nowotworowego, co może wspierać planowanie terapii i poprawiać efektywność pracy klinicznej. Wskazano jednak na konieczność dalszej poprawy w rozpoznawaniu zmian o małej intensywności metabolicznej i objętości, a także zwiększenia różnorodności danych treningowych.

Badanie Zang i wsp. [16] oceniło zastosowanie DL na obrazach 68Ga-PSMA-11 PET/CT w przewidywaniu patologicznego zaawansowania u pacjentów z PCa. W analizie porównano modele 3D ResNet-18 i DenseNet-121, które w testach osiągnęły AUC odpowiednio 0,612 i 0,559. Model 3D ResNet-18 osiągnął dokładność 68,2%, czułość 54,6% i swoistość 81,8%. Połączenie danych klinicznych z sygnaturą modelu DL poprawiło wyniki, osiągając najwyższą dokładność 72,7% oraz swoistość 90,9%, choć czułość pozostała na poziomie 54,6%. Wyniki sugerują, że techniki DL mogą wspierać prognozowanie patologicznego zaawansowania, jednak wymagają dalszej walidacji na większych i zróżnicowanych zestawach danych.

W retrospektywnym badaniu Janbain i wsp. [17] oceniono zastosowanie modelu Losowego Lasu Przetrwania (ang. Random Survival Forest, RSF) oparte-go na obrazach PSMA-PET w przewidywaniu odpowiedzi biochemicznej u 1029 pacjentów z nawrotem PCa po radykalnej prostatektomii, poddanych radioterapii ratunkowej. Model RSF osiągnął wskaźniki Harrella (C-index) wynoszące 0,79 dla treningu, 0,72 dla walidacji wewnętrznej i 0,69 dla walidacji zewnętrznej, przewyższając wcześniejszy nomogram (C-index: 0,67–0,71). Skorygowane wyniki Brier Score wskazywały na spójność modelu (0,13–0,14). Kluczowymi predyktorami były poziomy PSA przed sRT i występowanie nawrotów w węzłach chłonnych na PSMA-PET. Wyniki sugerują, że RSF może poprawić dokładność prognozowania, wspierając decyzje terapeutyczne i personalizując leczenie w nawrotach biochemicznych PCa.

Badanie Cao i wsp. [18] oceniło zastosowanie radiomików PSMA-PET oraz parametrów klinicznych do predykcji 2-letniego przeżycia bez przerzutów (ang. Metastasis-Free Survival, MFS) u 92 pacjentów z oligometastatycznym, kastracyjno-czułym rakiem prostaty (ang. Oligometastatic Castration-Sensitive Prostate Cancer, omCSPC) po terapii ukierunkowanej na przerzuty (ang. Metastasis-Directed Therapy, MDT). Model Lasu Losowego (ang. Random Forest, RF) osiągnął dokładność 80% oraz AUC 0,82 w teście Leave-One-Out. W walidacji międzyinstytucjonalnej, model przewidział zdarzenia MFS z dokładnością 75% dla danych z JHH (AUC 0,77) i 78% dla danych z BU (AUC 0,80). Kluczowymi cechami były: Total Energy, Entropia i Standard Deviation z GTV w pre-PSMA-PET oraz Total Energy i Kontrast z post-PSMA-PET. Wyniki potwierdzają potencjał PSMA-PET i radiomików w prognozowaniu wyników terapii MDT, co może wspierać decyzje terapeutyczne i poprawiać opiekę nad pacjentami.

Badanie Wang i wsp. [19] porównujące skuteczność diagnostyczną całego ciała [68Ga]Ga-PSMA-11 PET/CT i krótkiego pola widzenia (ang. Small-Area Field of View, SAFOV) u pacjentów z wczesnym nawrotem PCa i poziomem PSA < 0,2 ng/mL wykazało wyższy wskaźnik wykrywania dla total-body PET/CT (73,75% vs. 43,75%, $p < 0,001$). Total-body PET/CT poprawiło również obrazowanie kości (31,25% vs. 13,75%, $p = 0,008$) oraz zmieniło schematy leczenia u 80% pacjentów, w porównaniu do 55% dla SAFOV PET/CT ($p = 0,001$). Wartości AUC i specyficzność były istotnie wyższe dla total-body PET/CT, co wskazuje na większą dokładność diagnostyczną, szczególnie w przypadku niskich poziomów PSA. Wyniki te podkreślają znaczenie całego ciała PET/CT

w bardziej precyzyjnym prowadzeniu terapii w przypadku wczesnych nawrotów PCa.

Badanie Li i wsp. [20] zaprezentowało automatyczne podejście oparte na DL do segmentacji i klasyfikacji wzorców wychwytu w PSMA PET/CT u pacjentów z nawrotowym i/lub przerzutowym PCa. Wewnętrzne testy wykazały skuteczność segmentacji podejścia DL z Dice 0,700, IoU 0,566, precyzją 0,809 i czułością 0,660. W testach zewnętrznych uzyskano Dice 0,680, IoU 0,548, precyzję 0,749 i czułość 0,740. Klasyfikacja wzorców wychwytu podejściem multi-modalnym osiągnęła dokładność 76,4% (AUC 0,863) na danych wewnętrznych i 79,6% (AUC 0,851) na zewnętrznych. Framework DL przewyższył inne podejścia, skutecznie różnicując zmiany podejrzane od niepodejrzanych o raka, co wspiera diagnozę i prognozowanie u pacjentów z zaawansowanym PCa.

Badanie Glemser i wsp. [11] oceniło zastosowanie [18F]-PSMA-1007 PET/MRI wzbogaconego o analizę semi-ilościową (ang. Dynamic Contrast-Enhanced Imaging, DCE) i AI w diagnostyce pierwotnego PCa. Analiza DCE pozwoliła na różnicowanie tkanek łagodnych od złośliwych za pomocą parametrów takich jak wash-in slope oraz fitted maximum contrast ratio, które w całej kohorcie skutecznie rozróżniały zmiany podejrzane od zdrowych. Z kolei pipeline AI oparty na DL automatycznie wykrył klinicznie istotne zmiany w 6/7 pacjentów z Gleason Score $\geq 7a$, osiągając zgodność DICE na poziomie 0,44 między manualną a automatyczną segmentacją. Połączenie tych metod zwiększyło dokładność diagnostyczną, umożliwiając precyzyjne lokalizowanie i klasyfikację zmian w stagingu PCa.

Przegląd systematyczny Liu i wsp. [21] ocenił zastosowanie AI w analizie PSMA PET w diagnostyce PCa z przerzutami oraz zajęciem węzłów chłonnych. AI osiągnęła wysoką dokładność (AUC do 98%) i czułość (62–97%) w wykrywaniu zmian przerzutowych, przy zmiennych wartościach PPV (39,2–66,8%). Modele AI były skuteczne w automatyzacji zadań, takich jak segmentacja organów i alokacja anatomiczna zmian, oraz w redukcji zmienności między obserwatorami. Wykorzystywano różne algorytmy, m.in. CNN i ML, z których jeden model osiągnął czułość 91,5% dla węzłów chłonnych regionalnych (aPROMISE). Modele AI różnicowały także zmiany przerzutowe od łagodnych (AUC 0,98, czułość 94%, specyficzność 89%) oraz przewidywały odpowiedź na terapię z czułością 75% i AUC 80%. Chociaż wyniki są obiecujące, potrzebne są większe, prospektywne badania z bardziej reprezentatywnymi zestawami danych przed pełną integracją AI w praktyce klinicznej.

W badaniu Holzschuh i wsp. [22] oceniono 161 pacjentów z pierwotnym PCa, z danych pochodzących z czterech ośrodków (Freiburg: $n = 96$, Monachium: $n = 19$, Cypr: $n = 32$, Drezno: $n = 14$). Wykorzystano modele CNN oparte na architekturze nnU-Net do automatycznego konturowania wewnątrzprostatycznych GTV na obrazach PET z 18F-PSMA-1007. Modele trenowano na danych wieloośrodkowych i jednookręgowych, a skuteczność oceniano za pomocą DSC oraz Hausdorff Distance (HD95). Model wieloośrodkowy osiągnął medianę DSC 0,76 (IQR: 0,64–0,84) przy cross-walidacji i 0,74 (IQR: 0,56–0,86) w podejściu leave-one-center-out ($p = 0,18$). W analizie per ośrodek, istotna poprawa wyników odnotowana została dla Drezna (DSC 0,71 vs. 0,68, $p < 0,001$) i Cypru (0,74 vs. 0,72, $p < 0,01$), podczas gdy dla Monachium (0,74 vs. 0,72, $p > 0,05$) i Freiburga (0,78 vs. 0,71, $p > 0,05$) różnice nie były statystycznie istotne. Wartości HD95 dla modeli wieloośrodkowych nie różniły się istotnie od modeli jednookręgowych w poszczególnych ośrodkach ($p > 0,05$). Badanie potwierdza, że trening na danych wieloośrodkowych zwiększa zdolność generalizacji modeli CNN w segmentacji GTV przy użyciu PET z 18F-PSMA-1007.

W retrospektywnym badaniu Urso i wsp. [23] 29 pacjentów z oligorekurencyjnym PCa oceniono modele ML bazujące na parametrach klinicznych oraz radiomicznych wyodrębnionych z obrazów 18F-cholina PET/CT. Modele miały na celu predykcję biochemicznej wznowy (ang. Biochemical Recurrence, BCR) po MDT. Z 46 zmian 76,1% dotyczyło grupy BCR1. W 2-letniej obserwacji BCR wystąpiło u 69% pacjentów. Spośród pięciu algorytmów ML, model Stochastic Gradient Descent (SGD) dla PET osiągnął najlepsze wyniki: AUC 0,95, CA 0,90, czułość 0,90, PPV 0,94, NPV 0,93. Model RF dla CT wykazał również wysoką skuteczność (AUC 0,92, CA 0,85, czułość 0,85, PPV 0,91, NPV 0,74). Modele PET przewyższyły CT w predykcji BCR, co przypisano lepszym możliwościom obrazowania metabolicznego 18F-cholina. Wskazano na konieczność walidacji na większych, zewnętrznych zbiorach danych, aby wprowadzić modele do codziennej praktyki klinicznej w kontekście medycyny spersonalizowanej.

2. Innowacje w diagnostyce i leczeniu raka płuc

Rak płuc (ang. Lung Cancer, LC) jest jedną z głównych przyczyn zgonów związanych z nowotworami na świecie, z 5-letnim wskaźnikiem przeżycia poniżej 20%. Nowoczesne terapie, takie jak immunoterapia i leczenie ukierunkowane molekularnie, znacząco poprawiają rokowania u pacjentów z mutacjami kierunkowymi, ale pozostają nieskuteczne u większości pacjentów. Medycyna

nuklearna z wykorzystaniem 18F-FDG PET/CT odgrywa kluczową rolę w diagnostyce, ocenie odpowiedzi na leczenie i monitorowaniu nawrotów. Włączenie AI w analizę obrazów pozwala na poprawę dokładności diagnostycznej i obniżenie błędów wynikających z subiektywnej oceny. AI umożliwia identyfikację wzorców niedostrzegalnych dla ludzkiego oka, co wspiera personalizację terapii, szczególnie w rozróżnianiu zmian związanych z leczeniem od nawrotu czy ocenie biomarkerów, takich jak PD-L1. Dzięki AI możliwe jest także zmniejszenie kosztów i czasu potrzebnego na diagnostykę oraz poprawa opieki nad pacjentami z LC [24-27].

W badaniu Yuan i wsp. [28] zastosowanie ML i DL na obrazach PET/CT znacząco poprawiło diagnostykę i prognozowanie LC. Algorytmy CNN wykryły 90% przypadków z poziomem automatycznej segmentacji porównywalnym do manualnej analizy, redukując fałszywie dodatnie wyniki z 72,8 do 4,9 na przypadek. Modele SVM osiągnęły czułość 95,6% i niski odsetek wyników fałszywie dodatnich. W prognozowaniu dynamiki zmian nowotworowych dla guzów powyżej 5 mL uzyskano AUC = 0,854. Dodatkowo ML umożliwiło diagnozę przy ultraniskich dawkach PET/CT (3,3%), osiągając czułość 91,5% i swoistość 94,2%. Modele ResNet-18 poprawiły dokładność identyfikacji mutacji EGFR z AUC w zakresie 0,81–0,86. Wnioski wskazują, że ML i DL w analizie PET/CT znacząco podnoszą dokładność diagnostyczną, redukują fałszywe wyniki i wspierają personalizację terapii przy niższym narażeniu na promieniowanie.

W badaniu Bo Li i wsp. [29] opracowano model radiomiki DL oparty na PET/CT w celu przewidywania ekspresji ligandu programowanej śmierci komórkowej (ang. Programmed Death-Ligand 1, PD-L1) u pacjentów z niedrobnokomórkowym rakiem płuca (ang. Non-Small Cell Lung Cancer, NSCLC). W analizie wzięło udział 136 pacjentów (42 PD-L1-negatywnych i 94 PD-L1-pozytywnych). Wyodrębniono łącznie 2446 cech radiomicznych i 4096 cech DL na pacjenta. W treningu model DL osiągnął AUC = 0,935 (95% CI: 0,865–0,975), a model radiomiczny AUC = 0,829 (95% CI: 0,738–0,898). Po fuzji obu modeli (Nomo_model) uzyskano AUC = 0,954 (95% CI: 0,890–0,986) w kohorcie treningowej i 0,910 (95% CI: 0,779–0,977) w walidacyjnej. Czułość i swoistość Nomo_model wynosiły odpowiednio 80,88% i 96,30% w treningu oraz 76,92% i 100% w walidacji. Model ten oferuje precyzyjne i nieinwazyjne narzędzie wspierające identyfikację pacjentów PD-L1-pozytywnych, poprawiając decyzje terapeutyczne.

W badaniu Zhang i wsp. [30] opracowano modele oparte na radiomice i cechach kliniczno-metabolicznych przy użyciu PET/CT z 18F-FDG w celu

różnicowania podtypów histologicznych LC: raka płaskonabłonkowego (ang. Squamous Cell Carcinoma, SCC) i gruczolakoraka. W analizie wzięło udział 255 pacjentów z NSCLC, podzielonych na kohorty treningowe ($n=177$) i walidacyjne ($n=78$). Zidentyfikowano cechy kliniczne (płeć, wiek, CEA [Carcino-embryonic Antigen, Antygen rakowo-płodowy], SUVmax, SUVmin [Wartość standaryzowanego wychwytu, Standardized Uptake Value]) oraz radiomiczne, które w połączeniu dały model z najwyższym AUC = 0,870 (95% CI: 0,791–0,948) w walidacji. Model radiomiczny osiągnął AUC = 0,774, a kliniczny AUC = 0,848. Modele testowano przy użyciu algorytmu regresji logistycznej, który zapewnił najwyższą czułość i swoistość w walidacji. Połączenie radiomiki i cech klinicznych poprawiło skuteczność przewidywań, oferując narzędzie wspierające decyzje kliniczne w diagnostyce NSCLC.

W badaniu Ju i wsp. [31] opracowano system stratyfikacji ryzyka (ang. Clinical Risk Stratification, CRS) oparty na cechach DL i całkowitej metabolicznej objętości guza w obrazach PET/CT u pacjentów z NSCLC. Zbadano dane 590 pacjentów (413 w kohorcie treningowej i 177 w walidacyjnej). CRS osiągnął lepsze wskaźniki w porównaniu z klasyfikacją TNM (międzynarodowy system klasyfikacji nowotworów, Tumor-Nodes-Metastasis) w prognozowaniu całkowitego przeżycia (ang. Overall Survival, OS) i przeżycia wolnego od progresji (ang. Progression-Free Survival, PFS), osiągając C-index dla OS = 0,71 (trening) i 0,73 (walidacja) oraz dla PFS = 0,702 i 0,732. Połączenie CRS z TNM w nomogramie zwiększyło skuteczność predykcijną, osiągając C-index = 0,741 (trening) i 0,771 (walidacja). CRS umożliwił dokładniejszą identyfikację pacjentów wysokiego ryzyka, przewidyując 9-krotnie wyższe ryzyko śmierci w najwyższym poziomie ryzyka w porównaniu z najniższym. Narzędzie to wspiera lekarzy w podejmowaniu decyzji terapeutycznych i poprawia prognozy pacjentów z NSCLC.

W badaniu retrospektywnym Torbati i wsp. [32] oceniono 79 pacjentów z zaawansowanym NSCLC (54 mężczyzn, 25 kobiet), u których przed terapią wykonano badanie [18F]FDG PET/CT. Pacjentów podzielono na kohorty treningową ($n=44$) i testową ($n=35$). Ekstrakcja cech teksturalnych z pierwotnych guzów i przerzutów w węzłach chłonnych wykonano za pomocą programu LIFEx. Spośród sześciu metod ML, najwyższą dokładność w predykcji odległych przerzutów uzyskano dla drzewa decyzyjnego (dokładność 74,4%, AUC 0,63). Optymalna kombinacja cech (shape sphericity, GLRLM_RLNU, TLG [Całkowita glikoliza zmiany], MTV [Metaboliczna objętość guza], shape compacity) pozwalała na przewidywanie ryzyka przerzutów z akceptowalną skutecznością.

Wyniki sugerują, że analiza teksturalna połączona z klasycznymi metodami ML może wspierać decyzje terapeutyczne u pacjentów z wysokim ryzykiem przerzutów. Limitacje badania obejmują retrospektywny charakter, niewielką liczbę pacjentów oraz brak walidacji zewnętrznej.

W retrospektywnym badaniu Hosseini i wsp. [33] z udziałem 126 pacjentów z NSCLC (średni wiek 47 ± 12 lat; 28,6% z LVI) oceniano różne metody segmentacji obrazów PET ([^{18}F]FDG), selekcji cech radiomicznych i algorytmy ML w predykcji inwazji limfowaskularnej (ang. Lymphovascular Invasion, LVI). Spośród wielu kombinacji najlepsze wyniki osiągnięto przy zastosowaniu metody segmentacji Iterative Thresholding (IT) z progami 45%, algorytmu Recursive Feature Elimination (RFE) oraz klasyfikatora regresji logistycznej (LR) wspomaganego przez SMOTE (AUC = 0,93, dokładność = 0,84, czułość = 0,85, swoistość = 0,84). Alternatywnie, metoda segmentacji Fuzzy-C-Mean z selekcją MRMR i klasyfikatorem LR osiągnęła AUC = 0,92 i dokładność = 0,87. Wyniki wskazują, że wybór odpowiedniej kombinacji segmentacji i algorytmów ML umożliwia skuteczną predykcję LVI w NSCLC. Limitacje badania obejmowały niewielką liczebność próby i brak walidacji zewnętrznej, co podkreśla potrzebę dalszych badań z większymi, wieloośrodkowymi zbiorami danych.

W retrospektywnym badaniu Zhu i wsp. [34] oceniano 226 pacjentów z NSCLC (średni wiek 58,7 lat; 34% z przerzutami do mózgu) pod kątem predykcji przerzutów do mózgu (ang. Brain Metastases, BM) za pomocą danych radiomicznych z obrazów PET/CT ([^{18}F]FDG). Analizowano cechy płytkie i głębokie, wykorzystując PyRadiomics oraz 3D NN (C3D, Res18, R3D). Najwyższą skuteczność w predykcji BM osiągnął model C3D_LungBrain, integrujący cechy płuc i mózgu (AUC = 0,964, dokładność = 95%, F1 = 96%). Modele oparte wyłącznie na cechach płuc (ONLY_TRAD, AUC = 0,531) były znacząco mniej efektywne. Wykazano stabilność cech głębokich niezależnie od NN (AUC > 0,95). Wyniki sugerują, że integracja cech płytkich i głębokich z płuc i mózgu poprawia predykcję BM, co może wspierać personalizację leczenia pacjentów z NSCLC. Limitacje badania obejmowały brak zewnętrznej walidacji i trudność interpretacji cech głębokich w kontekście klinicznym.

W retrospektywnym badaniu Ortiz i wsp. [24] analizowano dane 309 pacjentów z potwierdzonym LC (103 w każdej grupie: gruczolakorak, rak płaskonabłonkowy, drobnokomórkowy) za pomocą obrazów PET/CT i metod ML. Stosowano modele Stacking i Boosting (AdaBoost) w celu oceny ich skuteczności w wykrywaniu LC. Model Stacking osiągnął najwyższą dokładność (94%), precyzję (94%), czułość (94%) i F1-score (94%) w porównaniu do modelu Boosting,

który uzyskał niższe wyniki we wszystkich kategoriach (77%). Stosowanie technik równoważenia danych pozwoliło na uniknięcie problemów z nadmiernym dopasowaniem. Wyniki podkreślają potencjał złożonych modeli, takich jak Stacking, w integracji różnych algorytmów bazowych (RF, SVM, regresja logistyczna) dla poprawy jakości predykcji w diagnostyce LC. Limity badania obejmują retrospektywny charakter i niewielką liczbę pacjentów, co wymaga dalszych badań prospektywnych.

W retrospektywnym badaniu Zhao i wsp. [25] oceniano skuteczność siedmiu modeli DL w klasyfikacji podtypów histologicznych NSCLC (adenocarcinoma, ADC; squamous cell carcinoma, SqCC) na podstawie obrazów PET/CT z 18F-FDG. Analizowano dane 189 pacjentów (107 ADC, 82 SqCC), z czego Mobilenet v2 osiągnął najlepsze wyniki ($AUC = 0,744$, $AP = 0,759$ w wewnętrznej kohorcie testowej; $AUC = 0,767$, $AP = 0,768$ w rozszerzonej kohorcie testowej). Model wykazał wysoką precyzję (0,750), czułość (0,818), specyficzność (0,688) i F1-score (0,800) w rozszerzonej kohorcie testowej. Wyniki wskazują na potencjał modelu Mobilenet v2 w nieinwazyjnej klasyfikacji NSCLC, z istotną wartością kliniczną dla pacjentów, którzy nie mogą być diagnozowani za pomocą biopsji. Limity obejmują niewielką próbę i retrospektywny charakter badania, co wymaga dalszych badań prospektywnych oraz większej liczby danych dla poprawy ogólności modelu.

W retrospektywnym badaniu Liang i wsp. [26] oceniano 101 pacjentów z NSCLC (średni wiek 64,6 lat, 57 mężczyzn, 44 kobiety), którzy przeszli badanie 18F-FDG PET/CT i analizę PD-L1. Wykorzystano DL z modelem DenseNet121 w trzech sesjach (PET, CT, PET+CT). Najwyższą skuteczność w predykcji ekspresji PD-L1 uzyskano dla modelu PET+CT połączonego z cechami klinicznymi ($AUC = 0,96 \pm 0,009$, dokładność = $0,95 \pm 0,025$, czułość = 0,958, swoistość = 0,962). Samo PET ($AUC = 0,814$) i CT ($AUC = 0,721$) miały niższą skuteczność. Połączenie danych obrazowych z cechami klinicznymi, takimi jak wiek, płeć, historia palenia, średnia SUV, maksymalna SUV i TLG, znacząco poprawiło dokładność predykcji. Wyniki potwierdzają potencjał 18F-FDG PET/CT jako biomarkera w przewidywaniu ekspresji PD-L1, z potencjalnym zastosowaniem w planowaniu terapii immunologicznej. Limity badania obejmują małą próbę oraz brak analizy warstwowej ekspresji PD-L1, co wymaga dalszych badań.

W retrospektywnym badaniu Sung i wsp. oceniono 308 pacjentów z rakiem płuca, którzy przeszli radioterapię (RT) i mieli podejrzenie nawrotu na podstawie obrazów 18F-FDG PET/CT. Wykorzystano 2D model CNN do odróżnienia zmian związanych z RT od nawrotu nowotworu. Model osiągnął

wysoką skuteczność diagnostyczną z AUC w zakresie 0,98–0,99, czułością 95–98% i swoistością 87–95% w pięciu niezależnych zestawach testowych. Porównanie wyników modelu z raportami formalnymi wykazało wyższą dokładność (95,5% vs. 85,4%) i swoistość (95,9% vs. 80,1%) modelu. Grad-CAM umożliwił wizualizację obszarów decyzyjnych modelu, co zwiększyło zrozumienie procesu klasyfikacji. Wyniki sugerują potencjał modelu jako narzędzia wspomagającego w diagnostyce nawrotów po RT, choć konieczne są dalsze badania wieloośrodkowe w celu walidacji zewnętrznej. Limitacje obejmowały brak zewnętrznej walidacji, ograniczoną liczebność próby oraz różnice w harmonizacji skanerów PET.

W prospektywnym badaniu Meng i wsp. [35] oceniono 145 pacjentów z NSCLC, w tym 101 w zbiorze treningowym (67 z przerzutami do węzłów chłonnych, 66,34%) i 44 w zbiorze testowym (29 z przerzutami, 65,91%). Opracowano model [18F]FDG PET/3D-UTE MRI oparty na czynnikach klinicznych, radiomice PET i 3D-UTE za pomocą metod ML. Model z wykorzystaniem RF osiągnął AUC 0,912 w zbiorze treningowym i 0,791 w testowym, przewyższając modele indywidualne (AUC dla PET: 0,828/0,768, 3D-UTE: 0,834/0,745, kliniczny: 0,838/0,756). Model PET/3D-UTE wykazał podobną skuteczność do modelu PET/CT (AUC: 0,890/0,793), przy czym różnice nie były istotne statystycznie. Czułość i swoistość modelu PET/3D-UTE wyniosły odpowiednio 72,41% i 86,67% w zbiorze testowym. Krzywe kalibracji wskazywały na dobrą zgodność (C-index: 0,912/0,791), a analiza DCA potwierdziła użyteczność kliniczną modelu. Model oferuje nieinwazyjną ocenę stanu węzłów chłonnych przy niższym obciążeniu promieniowaniem niż PET/CT.

3. Analiza obrazów, modelowanie prognostyczne i efektywność terapii

Współczesna medycyna nuklearna korzysta z zaawansowanych metod analizy obrazów i modelowania prognostycznego w celu oceny skuteczności terapii i prognozowania wyników leczenia. Integracja obrazów PET/CT z danymi klinicznymi, przy zastosowaniu AI i metod ML oraz DL, pozwala na kompleksową ocenę procesów metabolicznych i strukturalnych nowotworów oraz personalizację terapii. Techniki takie jak mechanizmy cross-attention, Q-Former czy projekcje MA-MIP umożliwiają efektywne wykorzystanie wielomodalnych danych, eliminując subiektywność manualnych segmentacji. Omówione w podsekcji metody podkreślają znaczenie integracji danych obrazowych i tekstowych w kontekście prognozowania przeżycia pacjentów i poprawy diagnostyki w onkologii [35-38].

W retrospektywnym badaniu Zhi i wsp. [39] przeanalizowano dane 327 pacjentów z rakiem żołądka poddanych obrazowaniu 18F-FDG PET/CT. Pacjentów losowo podzielono na zbiór treningowy ($n = 229$) i walidacyjny ($n = 98$). Z obrazów PET wyodrębniono 171 cech radiomicznych i skonstruowano wskaźniki PET RS za pomocą analizy LASSO i RSF. W modelu uwzględniono PET RS, wiek, poziom CEA i stadium TNM, co pozwoliło na przewidywanie OS. Model osiągnął C-index 0,817 (95% CI: 0,790–0,844) w zbiorze treningowym i 0,707 (95% CI: 0,640–0,774) w walidacyjnym. Pacjenci z wyższym PET RS mieli istotnie gorsze rokowanie (HR 7,99, $p < 0,001$). Model przewyższał skutecznością modele oparte wyłącznie na danych klinicznych (C-index: 0,630–0,730). Kalibracja i krzywe DCA wskazały na dobrą zgodność i użyteczność kliniczną modelu. Wyniki sugerują, że PET RS, w połączeniu z cechami klinicznymi, może poprawić personalizację terapii i prognozowanie w raku żołądka.

W retrospektywnym badaniu Jannusch i wsp. [40] obejmującym 143 pacjentki z rakiem piersi przeanalizowano dane kliniczne i obrazowe pochodzące z [18F]FDG-PET/MRI w celu predykcji całkowitej patologicznej odpowiedzi (ang. Pathological Complete Response, pCR) na neoadiuwantową terapię systemową. Do stworzenia modelu predykcyjnego wykorzystano 56 cech w modelu XGBoost (ML). Model osiągnął średnie AUC na poziomie 80,4% ($\pm 6,0\%$) w walidacji krzyżowej, przy czułości 54,5% ($\pm 21,3\%$), swoistości 83,6% ($\pm 4,2\%$), PPV 63,6% ($\pm 8,5\%$) oraz NPV 77,6% ($\pm 8,1\%$) dla progu decyzyjnego 0,5. Kluczowe cechy obejmowały dane histopatologiczne (ER+, Ki67, PR+, histologiczny stopień złośliwości) oraz wskaźniki PET, takie jak stosunek SUVmean zmiany do SUVmean aorty. Wyniki sugerują potencjał modelu w predykcji pCR, choć obecna wydajność nie jest wystarczająca do zastosowań klinicznych. Wskazano na potrzebę większych zbiorów danych i badań longitudinalnych dla poprawy skuteczności modelu.

W retrospektywnym badaniu Wu i wsp. [41] z udziałem 194 pacjentów z rakiem gardła oceniono predykcję wyników radioterapii na podstawie obrazów [18F]FDG PET/CT przy użyciu modeli DL. Modele trenowano na 162 pacjentach, a walidację przeprowadzono na 32 pacjentach. Używając metody Ensemble Voting Classifier, osiągnięto AUC dla lokalnej wznowy 0,850, nawrotu węzłów szyjnych 0,878 i odległych przerzutów 0,893. Dokładność wyniosła odpowiednio 87,5%, 68,8% i 78,1%. Najwyższą dokładność w predykcji lokalnej wznowy osiągnięto modelem ViT (AUC 0,831, dokładność 78,1%). Wyniki sugerują, że modele DL mogą wspierać predykcję odpowiedzi na radioterapię, ale predykcja nawrotu węzłów szyjnych i odległych przerzutów wymaga dalszej optymalizacji,

m.in. przez integrację parametrów klinicznych. Wskazano na konieczność walidacji zewnętrznej w celu potwierdzenia użyteczności klinicznej modeli.

W badaniu Duan i wsp. [42] z udziałem 23 pacjentów z miejscowo zaawansowanym NSCLC oceniono model GWO+RuleFit (ML), łączący interpretowalny algorytm RuleFit z heurystycznym algorytmem Gray Wolf Optimizer (GWO) w celu przewidywania odpowiedzi na chemioradioterapię na podstawie zmian SUV na obrazach FDG-PET w trakcie leczenia. W analizie klasyfikacyjnej osiągnięto AUC w zakresie 0,58–0,86, przewyższając inne modele w większości konfiguracji. W regresji osiągnięto RMSE w zakresie 0,162–0,192 dla niskowymiarowych modeli i 0,189–0,200 dla wysokowymiarowych modeli opartych na radiomice, co było istotnie lepsze dla modeli wysokowymiarowych w czterech z pięciu analizowanych regionów ($p < 0,004$). Model GWO+RuleFit wykazał wyższą interpretowalność, dzięki czemu może wspierać podejmowanie decyzji klinicznych w zakresie stratygrafii ryzyka i personalizacji terapii. Wskazano na potrzebę dalszych badań z większymi kohortami dla walidacji i ulepszenia modelu.

W retrospektywnym badaniu De Biase i wsp. [43] 399 pacjentów z rakiem jamy ustnej i gardła poddano ocenie modele predykcyjne DL, wykorzystując dane PET/CT oraz mapy prawdopodobieństwa guza (ang. Tumor Probability Map, TPM). Modele przewidywały lokalną kontrolę, kontrolę regionalną, przeżycie wolne od odległych przerzutów (DMFS) oraz przeżycie całkowite (OS). Mapy TPM zastąpiły tradycyjne maski objętości guza pierwotnego (GTVp). Najlepsze wyniki uzyskano dla lokalnej kontroli (C-index 0,74) i regionalnej kontroli (0,60) w modelach z TPM. Dla OS, modele z TPM i GTVp wykazały podobną skuteczność (C-index 0,82). Hybrydowe modele (łączące predykcje DL i dane kliniczne) osiągnęły najwyższe wartości C-index dla OS (0,82) i DMFS (0,81). Wyniki sugerują, że TPM dostarcza bardziej zróżnicowanych informacji niż GTVp, poprawiając predykcję wyników lokalnej kontroli i RC. Dalsze badania z udziałem większych zbiorów danych są konieczne dla walidacji zewnętrznej.

W retrospektywnym badaniu Kovacs i wsp. [37] oceniono skuteczność pięciu modeli DL do automatycznego wyznaczania GTV w obrazach 18F-FDG PET/CT u pacjentów z rakiem głowy i szyi. Wykorzystano 1095 wewnętrznych skanów do treningu ($n = 835$) i walidacji ($n = 260$) oraz 125 zewnętrznych skanów do testowania. Model nnU-Net osiągnął najwyższy współczynnik DSC 0,80 (95% CI: 0,77–0,86) w testach wewnętrznych i podobny wynik w testach zewnętrznych (DSC 0,81). Wewnętrzna zmienność między AI a ekspertami (DSC 0,78) była porównywalna do zmienności między ekspertami (DSC 0,82, $p = 0,12$). Analiza biomarkerów opartych na GTV wykazała dobre porozumienie

między AI a ekspertami (LoA dla SUVmean: -1,1 do 2,1 dla AI, -2,1 do 2,0 dla ekspertów). Wskazano, że DL może automatyzować wyznaczanie GTV w planowaniu radioterapii i poprawiać powtarzalność oraz stabilność biomarkerów. Badanie podkreśla potrzebę dalszych badań nad integracją AI w codziennej praktyce klinicznej, w tym szczególnie w skomplikowanych przypadkach, takich jak obrazy pooperacyjne.

W badaniu Toosi i wsp. [36] przeanalizowano dane 489 pacjentów z rakiem głowy i szyi zebrane z 7 ośrodków. Opracowano metodę bezsegmentacyjną opartą na głębokim uczeniu (DL) do predykcji RFS na podstawie obrazów FDG PET/CT. Zamiast manualnej segmentacji, wykorzystano automatyczne przycinanie regionu głowy i szyi oraz wielokątowe maksymalne projekcje intensywności (MA-MIP). Zastosowanie modelu EfficientNet Large z globalnym średnim pool-owaniem i fuzją cech metodą maksymalizacji kanałowej osiągnęło najwyższy c-index $0,719 \pm 0,032$ w walidacji krzyżowej na pięciu zestawach testowych, przewyższając wcześniejsze najlepsze metody z HECKTOR 2022 (c-index 0,688). Metoda wykazała wysoką powtarzalność i efektywność, eliminując subiektywność manualnej segmentacji i zwiększając wydajność procesu predykcji. Wyniki sugerują potencjał tej metody do zastosowań klinicznych przy ograniczonych zasobach obliczeniowych.

W retrospektywnym badaniu Wang i wsp. [38] na danych 642 pacjentów z rakiem głowy i szyi zebranych z siedmiu ośrodków zastosowano nowatorski model DL, łączący dane tekstowe (np. TNM, wiek, płeć) oraz obrazy PET/CT, wykorzystując mechanizmy cross-attention i Q-Former do integracji wielomodalnych danych. W predykcji OS, RFS, MFS oraz PFS model osiągnął najwyższe C-indeksy: 0,796 (MFS), 0,691 (OS), 0,641 (PFS) i 0,626 (RFS). Integracja PET i CT poprzez cross-attention przewyższyła wyniki uzyskane przy użyciu samych danych PET lub CT (C-indeksy: 0,730–0,742). Zastosowanie proponowanej funkcji strat (loss) znacząco poprawiło dokładność predykcji (C-indeks do 0,796), przewyższając tradycyjny model Coxa. Wyniki wskazują na potencjał multi-modalnych modeli DL w personalizacji medycyny, eliminując ograniczenia segmentacji i integrując dane tekstowe z obrazami.

W badaniu Wang i wsp. [44] opracowano nomogram oparty na cechach DL, radiomice oraz parametrach klinicznych w celu predykcji przerzutów do węzłów chłonnych w raku jelita grubego przy użyciu obrazów 18F-FDG PET/CT. W badaniu wzięło udział 119 pacjentów (54 z przerzutami i 65 bez), z kohortą walidacyjną 33 pacjentów. Model DLR Nomogram osiągnął AUC = 0,934 (95% CI: 0,884–0,983) w treningu, 0,902 (95% CI: 0,769–1,000) w walidacji i 0,836

(95% CI: 0,673–0,998) w testach. Model przewyższał klasyczne podejścia oparte na cechach radiomicznych (AUC = 0,876 w treningu) i klinicznych (AUC = 0,817 w treningu). Uwzględnienie parametrów takich jak SUVmax, krótkie średnice węzłów chłonnych oraz obszary tłuszczowe przyczyniło się do poprawy precyzji predykcji. Model ułatwia rozróżnienie między stanem zapalnym a przerzutami, co zwiększa jego przydatność kliniczną w planowaniu terapii i podejmowaniu decyzji terapeutycznych. DLR Nomogram oferuje zintegrowane podejście, które znacząco podnosi dokładność diagnostyczną w porównaniu z tradycyjnymi metodami obrazowania i klasyfikacji.

PODSUMOWANIE GŁÓWNYCH WYNIKÓW

Przegląd literatury wykazał, że AI znacząco wpływa na poprawę diagnostyki i personalizację terapii w medycynie nuklearnej. Modele AI osiągają wysoką precyzję diagnostyczną, m.in. AUC 0,98 w przewidywaniu progresji PCa oraz AUC 0,95 w diagnostyce LC. Automatyczna segmentacja obrazów pozwoliła na redukcję czasu analizy z godzin do minut, przy jednoczesnym zachowaniu wysokiej zgodności z ocenami ekspertów (Dice > 0,80). Wyniki potwierdzają potencjał AI w standaryzacji procesu diagnostycznego oraz zwiększaniu skuteczności decyzji klinicznych. Konieczne jest jednak dalsze badanie możliwości pełnej integracji AI w praktyce klinicznej.

DYSKUSJA

AI odgrywa coraz istotniejszą rolę w medycynie nuklearnej, rewolucjonizując procesy diagnostyczne i terapeutyczne poprzez automatyzację skomplikowanych zadań oraz identyfikację wzorców niewidocznych dla ludzkiego oka. Przykładem jest diagnostyka PCa, gdzie zastosowanie technik radiomiki i głębokiego uczenia pozwoliło osiągnąć AUC na poziomie 0,98. Podobne korzyści obserwowano w LC, gdzie algorytmy CNN zmniejszyły liczbę fałszywie dodatnich wyników oraz umożliwiły analizę obrazów o ultraniskich dawkach promieniowania. Modele predykcyjne integrujące dane kliniczne i obrazowe wspierają personalizację terapii, dostosowując leczenie do indywidualnych potrzeb pacjenta. AI zwiększa także powtarzalność wyników i standaryzację diagnostyki, co ma kluczowe znaczenie w wieloośrodkowych badaniach klinicznych.

OGRANICZENIA I WYZWANIA

Pomimo obiecujących wyników, wdrożenie AI w medycynie nuklearnej napotyka na liczne trudności. Brakuje wieloośrodkowych badań prospektywnych, co utrudnia walidację modeli na dużych populacjach. Ponadto różnorodność danych oraz brak standaryzacji protokołów obrazowania ograniczają możliwości generalizacji wyników. Kolejnym wyzwaniem jest interpretowalność algorytmów DL, co wpływa na zaufanie klinicystów. Dodatkowe bariery obejmują wysokie koszty implementacji, wymagania sprzętowe oraz konieczność ochrony danych zgodnie z regulacjami prawnymi. Aby zapewnić skuteczne wdrożenie AI w codziennej praktyce klinicznej, niezbędne są dalsze badania nad harmonizacją danych i usprawnieniem modeli.

PERSPEKTYWY ROZWOJU

Perspektywy rozwoju AI w medycynie nuklearnej obejmują zastosowanie uczenia federacyjnego, które umożliwia analizę danych z wielu ośrodków bez naruszania ich prywatności. Rozwój algorytmów generatywnych wspierających identyfikację nowych biomarkerów oraz integracja danych multimodalnych (obrazowanie, genomika, dane kliniczne) może zwiększyć precyzję diagnostyki i personalizacji terapii. Rozwiązania takie jak Grad-CAM poprawiają interpretowalność algorytmów, zwiększając zaufanie klinicystów. Ponadto optymalizacja zasobów obliczeniowych i redukcja kosztów implementacji mogą ułatwić szersze wdrożenie AI w różnych ośrodkach, przyczyniając się do rewolucji w medycynie nuklearnej.

WNIOSKI

Sztuczna inteligencja wykazuje ogromny potencjał w medycynie nuklearnej, znacząco poprawiając precyzję diagnostyki oraz skuteczność personalizacji terapii. Automatyzacja procesów, takich jak segmentacja obrazów, zmniejsza czas analizy i obciążenie personelu medycznego. Pomimo obiecujących wyników, pełne wdrożenie AI wymaga dalszych badań nad walidacją modeli, poprawą ich interpretowalności oraz harmonizacją danych. Pełna integracja AI z systemami opieki zdrowotnej ma potencjał do rewolucjonizacji diagnostyki i terapii, oferując nową jakość w opiece nad pacjentem.

REFERENCJE

1. Mintz Y, Brodie R. Introduction to artificial intelligence in medicine. *Minimally Invasive Therapy & Allied Technologies*. 2019;28(2):73-81. doi:<https://doi.org/10.1080/13645706.2019.1575882>
2. Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism*. 2017;69(69):S36-S40. doi:<https://doi.org/10.1016/j.metabol.2017.01.011>
3. Nensa F, Demircioglu A, Rischpler C. Artificial Intelligence in Nuclear Medicine. *Journal of Nuclear Medicine*. 2019;60(Supplement 2):29S-37S. doi:<https://doi.org/10.2967/jnumed.118.220590>
4. Balyen L, Peto T. Promising Artificial Intelligence-Machine Learning-Deep Learning Algorithms in Ophthalmology. *Asia-Pacific Journal of Ophthalmology*. 2019;8(3). doi:<https://doi.org/10.22608/apo.2018479>
5. Tafti D, Banks KP. Nuclear Medicine Physics. StatPearls. Published online Spring 2023. <https://pubmed.ncbi.nlm.nih.gov/33760490/>
6. Najam H, Dearborn MC, Tafti D. Nuclear Medicine Instrumentation. StatPearls. Published online Summer 2023. <https://pubmed.ncbi.nlm.nih.gov/37983326/>
7. Heston TF, Tafti D. Nuclear Medicine Safety. StatPearls. Published online Summer 2024. <https://pubmed.ncbi.nlm.nih.gov/38753905/>
8. Eli A, Oladele S, Ping X, Shai L. Detecting Fraudulent Claims in Nuclear Medicine Using Machine Learning Algorithms. ResearchGate. Published November 19, 2024. <https://www.researchgate.net/publication/385943265>
9. Xu K, Kang H. A Review of Machine Learning Approaches for Brain Positron Emission Tomography Data Analysis. *Nuclear Medicine and Molecular Imaging*. 2024;58(4):203-212. doi:<https://doi.org/10.1007/s13139-024-00845-6>
10. Robert JH, Slomka PJ. Artificial Intelligence in Nuclear Cardiology: An Update and Future Trends. *Seminars in Nuclear Medicine*. 2024;54(5). doi:<https://doi.org/10.1053/j.semnuclmed.2024.02.005>

11. Glemser PA, Freitag M, Kovacs B, et al. Enhancing the diagnostic capacity of [18F]PSMA-1007 PET/MRI in primary prostate cancer staging with artificial intelligence and semi-quantitative DCE: an exploratory study. Deleted Journal. 2024;8(1). doi:<https://doi.org/10.1186/s41824-024-00225-5>
12. Liu J, Cundy TP, Dixon, Desai N, Marimuthu Palaniswami, Lawrentschuk N. A systematic review on artificial intelligence evaluating PSMA PET scan for intraprostatic cancer. BJU International. 2024;134(5). doi:<https://doi.org/10.1111/bju.16412>
13. Tapper W, Carneiro G, Christos Mikropoulos, Thomas SA, Evans PM, Stergios Boussios. The Application of Radiomics and AI to Molecular Imaging for Prostate Cancer. Journal of Personalized Medicine. 2024;14(3):287. doi:<https://doi.org/10.3390/jpm14030287>
14. Sarah Lindgren Belal, Frantz S, Minarik D, et al. Applications of Artificial Intelligence in PSMA PET/CT for Prostate Cancer Imaging. Seminars in Nuclear Medicine. 2023;54(1). doi:<https://doi.org/10.1053/j.semnuclmed.2023.06.001>
15. Huang B, Yang Q, Li X, et al. Deep learning-based whole-body characterization of prostate cancer lesions on [68Ga]Ga-PSMA-11 PET/CT in patients with post-prostatectomy recurrence. European Journal of Nuclear Medicine and Molecular Imaging. 2023;51(4):1173-1184. doi:<https://doi.org/10.1007/s00259-023-06551-3>
16. Zang S, Jiang C, Zhang L, et al. Deep learning based on 68Ga-PSMA-11 PET/CT for predicting pathological upgrading in patients with prostate cancer. Frontiers in Oncology. 2024;13. doi:<https://doi.org/10.3389/fonc.2023.1273414>
17. Janbain A, Farolfi A, Armelle Guenegou-Arnoux, et al. A Machine Learning Approach for Predicting Biochemical Outcome After PSMA-PET-Guided Salvage Radiotherapy in Recurrent Prostate Cancer After Radical Prostatectomy: Retrospective Study. JMIR Cancer. 2024;10:e60323. doi:<https://doi.org/10.2196/60323>
18. Cao Y, Sutura P, Silva Mendes W, et al. Machine learning predicts conventional imaging metastasis-free survival (MFS) for oligometastatic castration-sensitive prostate cancer (omCSPC) using prostate-specific

- membrane antigen (PSMA) PET radiomics. *Radiotherapy and Oncology*. 2024;199:110443. doi:<https://doi.org/10.1016/j.radonc.2024.110443>
19. Wang Y, Dong L, Zhao H, et al. The superior detection rate of total-body [68Ga]Ga-PSMA-11 PET/CT compared to short axial field-of-view [68Ga]Ga-PSMA-11 PET/CT for early recurrent prostate cancer patients with PSA < 0.2 ng/mL after radical prostatectomy. *European Journal of Nuclear Medicine and Molecular Imaging*. 2024;51(8):2484-2494. doi:<https://doi.org/10.1007/s00259-024-06674-1>
 20. Li Y, Imami MR, Zhao L, et al. An Automated Deep Learning-Based Framework for Uptake Segmentation and Classification on PSMA PET/CT Imaging of Patients with Prostate Cancer. *Deleted Journal*. 2024;37(5):2206-2215. doi:<https://doi.org/10.1007/s10278-024-01104-y>
 21. Liu J, Cundy TP, Dixon, Lawrentschuk N. A Systematic Review on Artificial Intelligence Evaluating Metastatic Prostatic Cancer and Lymph Nodes on PSMA PET Scans. *Cancers*. 2024;16(3):486. doi:<https://doi.org/10.3390/cancers16030486>
 22. Holzschuh JC, Mix M, Freitag MT, et al. The impact of multicentric datasets for the automated tumor delineation in primary prostate cancer using convolutional neural networks on 18F-PSMA-1007 PET. *Radiation Oncology*. 2024;19(1). doi:<https://doi.org/10.1186/s13014-024-02491-w>
 23. Urso L, Corrado Cittanti, Manco L, et al. ML Models Built Using Clinical Parameters and Radiomic Features Extracted from 18F-Choline PET/CT for the Prediction of Biochemical Recurrence after Metastasis-Directed Therapy in Patients with Oligometastatic Prostate Cancer. *Diagnostics*. 2024;14(12):1264. doi:<https://doi.org/10.3390/diagnostics14121264>
 24. Guzmán Ortiz S, Hurtado Ortiz R, Jara Gavilanes A, Ávila Faican R, Parra Zambrano B. A serial image analysis architecture with positron emission tomography using machine learning combined for the detection of lung cancer. *Revista Española de Medicina Nuclear e Imagen Molecular (English Edition)*. 2024;43(3):500003. doi:<https://doi.org/10.1016/j.remnie.2024.5000034>

25. Zhao H, Su Y, Lyu Z, et al. Non-invasively Discriminating the Pathological Subtypes of Non-small Cell Lung Cancer with Pretreatment 18 F-FDG PET/CT Using Deep Learning. *Academic Radiology*. 2024;31(1):35-45. doi:<https://doi.org/10.1016/j.acra.2023.03.032>
26. Liang C, Zheng M, Zou H, et al. Deep learning-based image analysis predicts PD-L1 status from 18F-FDG PET/CT images in non-small-cell lung cancer. *Frontiers in Oncology*. 2024;14. doi:<https://doi.org/10.3389/fonc.2024.1402994>
27. Sung C, Oh JS, Park BS, Kim SS, Song SY, Lee JJ. Diagnostic performance of a deep-learning model using 18F-FDG PET/CT for evaluating recurrence after radiation therapy in patients with lung cancer. *Annals of Nuclear Medicine*. 2024;38(7):516-524. doi:<https://doi.org/10.1007/s12149-024-01925-5>
28. Yuan L, An L, Zhu Y, et al. Machine Learning in Diagnosis and Prognosis of Lung Cancer by PET-CT. *Cancer Management and Research*. 2024;Volume 16:361-375. doi:<https://doi.org/10.2147/cmar.s451871>
29. Li B, Su J, Liu K, Hu C. Deep learning radiomics model based on PET/CT predicts PD-L1 expression in non-small cell lung cancer. *European Journal of Radiology Open*. 2024;12:100549. doi:<https://doi.org/10.1016/j.ejro.2024.100549>
30. Zhang Y, Liu H, Chang C, Yin Y, Wang R. Machine learning for differentiating lung squamous cell cancer from adenocarcinoma using Clinical-Metabolic characteristics and 18F-FDG PET/CT radiomics. *PloS One*. 2024;19(4):e0300170. doi:<https://doi.org/10.1371/journal.pone.0300170>
31. Ju L, Li W, Zuo R, et al. Deep Learning Features and Metabolic Tumor Volume Based on PET/CT to Construct Risk Stratification in Non-small Cell Lung Cancer. *Academic Radiology*. 2024;31(11):4661-4675. doi:<https://doi.org/10.1016/j.acra.2024.04.036>
32. Hakkak Moghadam Torbati A, Pellegrino S, Fonti R, Morra R, Placido SD, Vecchio SD. Machine Learning and Texture Analysis of [18F]FDG PET/CT Images for the Prediction of Distant Metastases in Non-Small-Cell Lung Cancer Patients. *Biomedicines*. 2024;12(3):472. doi:<https://doi.org/10.3390/biomedicines12030472>

33. Hosseini SA, Ghasem Hajianfar, Pardis Ghaffarian, et al. PET radiomics-based lymphovascular invasion prediction in lung cancer using multiple segmentation and multi-machine learning algorithms. *Physical and Engineering Sciences in Medicine*. 2024;47:1613-1625. doi:<https://doi.org/10.1007/s13246-024-01475-0>
34. Zhu Y, Cong S, Zhang Q, et al. Multimodal radiomics-based methods using deep learning for prediction of brain metastasis in non-small cell lung cancer with 18F-FDG PET/CT images. *Biomedical Physics & Engineering Express*. 2024;10(6):065011. doi:<https://doi.org/10.1088/2057-1976/ad7595>
35. Meng N, Feng P, Yu X, et al. An [18F]FDG PET/3D-ultrashort echo time MRI-based radiomics model established by machine learning facilitates preoperative assessment of lymph node status in non-small cell lung cancer. *European Radiology*. 2023;34(1):318-329. doi:<https://doi.org/10.1007/s00330-023-09978-2>
36. Toosi A, Shiri I, Zaidi H, Rahmim A. Segmentation-Free Outcome Prediction from Head and Neck Cancer PET/CT Images: Deep Learning-Based Feature Extraction from Multi-Angle Maximum Intensity Projections (MA-MIPs). *Cancers*. 2024;16(14):2538. doi:<https://doi.org/10.3390/cancers16142538>
37. Kovacs DG, Ladefoged CN, Andersen KF, et al. Clinical Evaluation of Deep Learning for Tumor Delineation on 18F-FDG PET/CT of Head and Neck Cancer. *Journal of Nuclear Medicine*. 2024;65(4):623-629. doi:<https://doi.org/10.2967/jnumed.123.266574>
38. Wang Z, Zheng C, Han X, Chen W, Lu L. An Innovative and Efficient Diagnostic Prediction Flow for Head and Neck Cancer: A Deep Learning Approach for Multi-Modal Survival Analysis Prediction Based on Text and Multi-Center PET/CT Images. *Diagnostics*. 2024;14(4):448. doi:<https://doi.org/10.3390/diagnostics14040448>
39. Zhi H, Xiang Y, Chen C, et al. Development and validation of a machine learning-based 18F-fluorodeoxyglucose PET/CT radiomics signature for predicting gastric cancer survival. *Cancer Imaging*. 2024;24(1). doi:<https://doi.org/10.1186/s40644-024-00741-4>

40. Jannusch K, Dietzel F, Bruckmann NM, et al. Prediction of therapy response of breast cancer patients with machine learning based on clinical data and imaging data derived from breast [18F]FDG-PET/MRI. *European Journal of Nuclear Medicine and Molecular Imaging*. 2023;51(5):1451-1461. doi:<https://doi.org/10.1007/s00259-023-06513-9>
41. Wu KC, Chen SW, Chang RF, et al. Early prediction of radiotherapy outcomes in pharyngeal cancer using deep learning on baseline [18F] Fluorodeoxyglucose positron emission Tomography/Computed tomography. *European Journal of Radiology*. 2024;181:111811. doi:<https://doi.org/10.1016/j.ejrad.2024.111811>
42. Duan C, Liu Q, Wang J, et al. GWO+RuleFit: rule-based explainable machine-learning combined with heuristics to predict mid-treatment FDG PET response to chemoradiation for locally advanced non-small cell lung cancer. *Physics in Medicine & Biology*. 2024;69(15):155018. doi:<https://doi.org/10.1088/1361-6560/ad6118>
43. De Biase A, Ma B, Guo J, et al. Deep learning-based outcome prediction using PET/CT and automatically predicted probability maps of primary tumor in patients with oropharyngeal cancer. *Computer Methods and Programs in Biomedicine*. 2023;244:107939. doi:<https://doi.org/10.1016/j.cmpb.2023.107939>
44. Wang H, Zhang J, Li Y, et al. Deep-learning features based on F18 fluorodeoxyglucose positron emission tomography/computed tomography (18F-FDG PET/CT) to predict preoperative colorectal cancer lymph node metastasis. *Clinical Radiology*. 2024;79(9):e1152-e1158. doi:<https://doi.org/10.1016/j.crad.2024.05.017>