

Abstrakt

Celem pracy było opracowanie oraz ocena skuteczności narzędzia informatycznego opartego na skali CTCAE (Common Terminology Criteria for Adverse Events) służącego do automatycznego generowania alertów klinicznych na podstawie wyników badań laboratoryjnych. Przedstawiono proces implementacji algorytmu oraz jego integrację z systemem analizy danych medycznych. Przeprowadzono testy skuteczności detekcji zdarzeń niepożądanych klasyfikowanych zgodnie z poziomami CTCAE. Uzyskane wyniki wskazują na wysoką zgodność automatycznego systemu z ocenami manualnymi, co sugeruje możliwość praktycznego zastosowania w klinicznym monitorowaniu pacjentów.

Słowa kluczowe: CTCAE, zdarzenia niepożądane, automatyzacja, analiza wyników badań, alerty kliniczne

1. Wstęp

Tradycyjne systemy automatycznej detekcji działań niepożądanych związanych z leczeniem wykorzystują zazwyczaj dwa podejścia. Z jednej strony są to wyrażenia regularne (ang. *regular expressions*), które umożliwiają ekstrakcję istotnych informacji z dokumentacji medycznej. Jest to rozwiązanie stosunkowo łatwe do wdrożenia dla części parametrów skali CTCAE (Common Terminology Criteria for Adverse Events). Z drugiej strony coraz częściej stosuje się modele oparte na sieciach neuronowych — w tym duże modele językowe (LLM, ang. *Large Language Models*) — które umożliwiają analizę języka naturalnego i ocenę znacznie większego spektrum objawów i kontekstów klinicznych.

Metody oparte na LLM mają niewątpliwą przewagę z uwagi na ich elastyczność i zdolność do interpretacji złożonego języka, jednak ich zastosowanie wiąże się z dużym zapotrzebowaniem na moc obliczeniową. Możliwe są w tym zakresie dwa rozwiązania: (1) skorzystanie z usług chmurowych, co jednak wiąże się z ryzykiem nieautoryzowanego przetwarzania danych przez dostawcę platformy, oraz (2) lokalne przetwarzanie danych, które z kolei wymaga kosztownej infrastruktury obliczeniowej. Choć pojawia się coraz więcej lekkich modeli językowych zoptymalizowanych pod kątem energooszczędnych GPU, to nadal lokalna implementacja LLM stanowi wyzwanie budżetowe i infrastrukturalne.

Niezależnie od sposobu przetwarzania danych, kluczowym problemem pozostaje integracja systemów analizujących działania niepożądane z istniejącym oprogramowaniem do zarządzania dokumentacją medyczną. Różnorodność rozwiązań dostępnych na rynku sprawia, że ustandaryzowanie takiej integracji jest trudne.

Cel pracy

Celem niniejszej publikacji jest przedstawienie uproszczonego podejścia do monitorowania objawów niepożądanych zgodnie ze skalą CTCAE, które umożliwia zautomatyzowaną ocenę z minimalnym nakładem pracy po stronie lekarza oraz bez potrzeby głębokiej ingerencji w system zarządzania danymi medycznymi.

2. Materiały i metody

W niniejszym projekcie analizie poddano proces leczenia pacjentów z chłoniakiem Hodgkina. W pierwszym etapie przygotowano zestaw skal CTCAE wypełnianych przez lekarza na podstawie standardowej dokumentacji medycznej. Na ich podstawie wyodrębniono najczęściej występujące parametry, których obecność była możliwa do ustalenia w sposób rutynowy — bez konieczności dodatkowych pomiarów czy konsultacji specjalistycznych.

Choć nie można jednoznacznie stwierdzić, czy dane działania niepożądane wynikają wyłącznie z leczenia przyczynowego, leczenia wspomagającego czy też zmian stylu życia (np. hospitalizacji), to ich obecność może stanowić cenną informację o przebiegu terapii oraz stanie pacjenta. Wybrane parametry następnie oceniono pod kątem możliwości ich automatyzacji przy pomocy wyrażeń regularnych.

Większość z nich okazała się możliwa do analizy w oparciu o prosty skrypt JavaScript uruchamiany po stronie klienta. Takie podejście pozwala na implementację rozwiązania bez konieczności przesyłania danych poza infrastrukturę szpitala. W przypadku użycia systemu AMMS, wdrożenie sprowadza się do umieszczenia pliku JS na serwerze i załączenia go w aplikacji — co stanowi minimalną ingerencję w istniejące środowisko.

Dla celów testowych stworzono symulator działania AMMS oraz algorytm analizy działań niepożądanych. Dodatkowo opracowano narzędzie wspierane przez model Gemini (wersja 2.0 light), które uzyskuje dostęp poprzez API — klucz użytkownika wpisywany jest ręcznie w przeglądarce i nie jest zapisywany w pamięci trwałej.

Część parametrów wymaga odniesienia do stanu bazowego (baseline), dlatego zaproponowano zapisanie danych referencyjnych w polu komentarza pierwszej wizyty w systemie AMMS. W czasie sesji dane są dostępne wyłącznie w zmiennych JavaScript, bez zapisu w plikach ani bazach danych.

Wybór jednostki chorobowej

Jako jednostkę chorobową wybrano chłoniaka Hodgkina (HL), ze względu na jego względnie niską częstość występowania, zwłaszcza w populacji dorosłych, w porównaniu do innych nowotworów hematologicznych. Z tego względu dostępność danych — w szczególności dotyczących działań niepożądanych leczenia — jest ograniczona, a tym samym przydatne wydaje się uwzględnienie tej jednostki zamiast bardziej poznanej. Ponadto, HL cechuje się dużą zmiennością przebiegu klinicznego oraz różnorodnością schematów terapeutycznych, co czyni go dobrym kandydatem do budowy elastycznego algorytmu.

Faza 1: Wybór skali oceny działań niepożądanych

Pierwszym etapem badania było wybranie odpowiedniej skali do oceny działań niepożądanych leczenia. Przeanalizowano kilka skal używanych w hematologii, a następnie — poprzez przegląd literatury — sprawdzono ich zastosowanie w kontekście jednostek chorobowych zbliżonych do HL. Przegląd literatury był oparty o poszukiwania prac, wykorzystujących każdą ze skal w onkologii, i dzizinach pokrewnych, udokumentowane w badaniach naukowych użycie w innych okolicznościach nie było brane pod uwagę. Ostatecznie wybrano skalę **CTCAE** (Common Terminology Criteria for Adverse Events) jako najbardziej precyzyjną i jednocześnie najłatwiejszą do automatyzacji.

Dodatkowo rozważono wykorzystanie skali **PRO-CTCAE**, wypełnianej przez pacjenta, jako możliwe uzupełnienie oceny, jednak ze względu na złożoność badania w tym wariantcie postanowiono się skupić wyłącznie na CTCAE. Wybrano wersję 5 — jako ostatnią dobrze ugruntowaną dla tej skali.

Faza 2: Weryfikacja możliwych do oceny parametrów CTCAE

W drugiej fazie sprawdzono, które spośród ponad 700 parametrów CTCAE można realistycznie ocenić w warunkach standardowej dokumentacji medycznej. Przeanalizowano elektroniczną dokumentację pacjentów z HL w klinice hematologii. Dla każdego wpisu próbowano dopasować możliwy do oceny parametr CTCAE.

Odrzucano te parametry, które:

- wymagały jednoczesnej oceny więcej niż dwóch badań,
- wymagały pogłębionej analizy stanu pacjenta, niemożliwej do oceny na podstawie epikryzy i dokumentacji postępowania,
- były ocenialne na podstawie badań wykonanych tylko jednokrotnie w trakcie hospitalizacji.

Dane zbierano w arkuszu Excel, notując nazwę badania oraz informację, czy wynik się powtarzał (co warunkowało ocenę zmian względem stanu wyjściowego).

Faza 3: Projekt aplikacji webowej do oceny działań niepożądanych

Na podstawie wyników z fazy 2 zaprojektowano aplikację webową. Formularz aplikacji zawierał jedynie te parametry CTCAE, które przeszły pozytywnie wcześniejszą selekcję. Dodano również pola dla identyfikatora pacjenta, daty wizyty (z opcją anonimizacji i pseudonimizacji), oraz zastosowanego schematu leczenia.

Za pomocą formularza oceniono działania niepożądane u pięciu losowo wybranych pacjentów hospitalizowanych w latach 2023–2025. Dla parametrów wymagających oceny względem wartości bazowej (baseline), wykorzystano dane z pierwszej wizyty lub – jeśli była długa – z pierwszych dni przed rozpoczęciem leczenia.

Dla Danych z badań takich jak EKG przyjęto ocenianie na podstawie epikryzy, danych z wizyty oraz schematu leczenia — w celu określenia, np. konieczność interwencji lub ogólny stan pacjenta – wymaganych dla stopni 2 i 3 w CTCAE. Co ciekawe, żaden z pięciu pacjentów nie wykazał nawet pierwszego stopnia zaburzeń ocenianych na podstawie EKG. Częstość występowania poszczególnych działań nieporządkanych oceniano poprzez sumę wyników dla poszczególnych wizyt – przy czym, jeśli ocena w ctcae wynosiła, np. 3 – to liczyła się jako 3, gdyż w ten sposób oceniono jednocześnie nasilenie działań nieporządkanych.

Najczęściej wykrywane działania niepożądane pochodziły z wyników morfologii krwi oraz obserwacji pielęgniarskiej. Są to dane wygodne do analizy z użyciem wyrażeń regularnych (REGEX), ponieważ stanowią agregaty wielu pomiarów.

Faza 4: Symulacja systemu AMMS

Na podstawie doświadczeń z fazy 3 stworzono **symulator AMMS**, zawierający tabelę z wizytami pacjentów. Po wybraniu konkretnej wizyty wyświetla on listę przeprowadzonych badań, a następnie — po wskazaniu badania — jego treść w polu tekstowym. Dodano również pole na komentarz do wizyty, odpowiadający danym przechowywanym na serwerze – tu przechowywany w localStorage.

Wszystkie dane pacjentów zostały wygenerowane z użyciem LLM **Gemini 2.0**, na podstawie ręcznie zaprojektowanego szablonu, opartego na układzie danych AMMS. Dla bezpieczeństwa nie wykorzystano żadnych danych rzeczywistych pacjentów. Wygenerowane dane zapisano jako stały tekst w pliku HTML symulatora.

Dane do symulatora zostały wygenerowane przez ChatGPT na podstawie instrukcji – załączonych w **załącznik 1)**

Faza 5: Skrypt oceniający parametry CTCAE

Ostatnia faza polegała na stworzeniu niezależnego skryptu oceniającego parametry CTCAE na podstawie danych z symulatora. Aby zminimalizować ingerencję w interfejs AMMS, zaprojektowano **sterowanie za pomocą skrótów klawiszowych** (zgodnych z Web Keyboard API) i myszki:

- Shift + Alt + p, służy do włączania lub wyłączania niniejszej usługi, przy czym jest ona domyślnie włączona

Użytkownik najpierw klika myszką interesujący go element, a następnie jeśli usługa jest aktywna – pojawi się okienko sterowania, a jeśli są nieprawidłowości w wynikach – również powiadomienie o tym.

Dane nie zawierały informacji osobowych, co teoretycznie umożliwiało ich analizę przez zewnętrzne modele językowe. Jednak z uwagi na niepewność, czy metadane nie pozwoliłyby na deanonimizację, **zrezygnowano z tego podejścia**.

Ze względu na ograniczone zasoby nie zastosowano algorytmu usuwającego dane osobowe, ponieważ symulacja i tak ich nie zawierała.

Jeśli chodzi o algorytm oparty na REGEX — z uwagi na złożoną procedurę oceny bezpieczeństwa przy wdrożeniu — zdecydowano się na **proste wyrażenia regularne**, różniące się jedynie danymi poszukiwanego ciągu tekstowego, a nie strukturą. Ułatwia to późniejsze dodawanie nowych parametrów bez konieczności edycji wzorca.

Po wykryciu nieprawidłowości, użytkownikowi wyświetlane jest **okienko dialogowe** (o pozycjonowaniu typu fixed) z informacjami oraz przyciskiem do jego zamknięcia.

3. Wyniki

Faza 1 – Wybór skali do oceny toksyczności leczenia

W pierwszym etapie badania skupiono się na identyfikacji skali, która umożliwiłaby możliwie najpełniejszą, a jednocześnie wykonalną retrospektywną ocenę działań niepożądanych leczenia chłoniaka Hodgkina (HL). Wybór ten został poprzedzony analizą literatury przedmiotu, w której porównano dostępne skale oceny toksyczności terapii onkologicznych, ze szczególnym uwzględnieniem specyfiki leczenia hematoonkologicznego.

Na podstawie przeprowadzonego przeglądu uznano, że najbardziej adekwatną skalą do oceny toksyczności leczenia HL jest skala **CTCAE (Common Terminology Criteria for Adverse Events)**. Skala ta jest powszechnie stosowana w badaniach onkologicznych, a jej struktura umożliwia zarówno ocenę prospektywną, jak i retrospektywną – co czyni ją szczególnie użyteczną w kontekście analizy elektronicznej dokumentacji medycznej. Dodatkowym atutem skali CTCAE jest dostępność jej wersji 5 w formacie edytowalnym (np. .xlsx), co umożliwia konwersję do formatu .xml oraz implementację w środowiskach webowych, bez konieczności wykorzystania serwera.

Alternatywą pozostaje skala **CTC**, jednak CTCAE okazała się lepiej dopasowana do potrzeb badania, ze względu na swoją szczegółowość i szerokie zastosowanie w różnych jednostkach chorobowych, w tym hematologicznych. Inne skale, takie jak **ASTCT**, mają zastosowanie głównie w kontekście terapii komórkowych (np. CAR-T) i nie były adekwatne do oceny działań niepożądanych standardowych schematów chemioterapii w HL.

Dodatkowo, rozważono zastosowanie pomocniczych narzędzi, takich jak skala **MAT** do oceny nudności i wymiotów, oraz kwestionariusze jakości życia (np. **Q-TWiST**, **EORTC QLQ-C30**), jednak ich zastosowanie uznano za bardziej odpowiednie dla badań prospektywnych i nie włączono ich do niniejszego modelu retrospektywnego. Podobnie, skale służące do oceny toksyczności radioterapii, takie jak **LENT-SOMA** czy **RTOG-EORTC**, uznano za zbędne w kontekście całościowej oceny działań niepożądanych leczenia systemowego – ich funkcję może bowiem pełnić CTCAE.

Skala **WHO Toxicity Grading Scale**, choć niekiedy stosowana w starszych pracach, została uznana za trudniejszą w implementacji w środowisku elektronicznym w porównaniu do CTCAE i z tego powodu również została odrzucona.

W ramach fazy pierwszej oceniono również potencjalne zastosowanie formularzy **PRO-CTCAE** (wersji skali CTCAE raportowanej przez pacjenta), jednak ze względu na złożoność oraz ograniczenia wynikające z retrospektywnego charakteru badania, zdecydowano się skoncentrować wyłącznie na klasycznej wersji CTCAE, ocenianej przez lekarzy na podstawie dokumentacji klinicznej.

Skala **WHO Toxicity Grading Scale** cechuje się prostotą oraz intuicyjnością, co czyni ją łatwą do zastosowania również w retrospektywnej analizie danych z dokumentacji klinicznej [1]. Jest ona publicznie dostępna w formie szczegółowego dokumentu (PDF), co ułatwia wdrożenie w badaniach niekomercyjnych lub lokalnych [1]. Niemniej jednak, analiza literatury oraz dokumentacji klinicznej nie potwierdza jej powszechnego wykorzystania w onkologii – dziedzinie, która wymaga bardziej wyspecjalizowanych narzędzi oceny działań niepożądanych terapii [1].

W ramach opublikowanych badań i opisów przypadków klinicznych, skala CTCAE została wykorzystana w kilku kontekstach klinicznych, co potwierdza jej szerokie zastosowanie i potrzebę dalszego jej rozwoju w odniesieniu do nowoczesnych terapii onkologicznych, w tym terapii celowanych, immunoterapii i terapii komórkowych. Poniżej skrótowy opis tych badań.

Reakcje na wlew brentuksymabu vedotin[2]

W przypadku 30-letniej pacjentki z nawrotowym chłoniakiem Hodgkina obserwowano ciężkie reakcje infuzyjne (nudności, ból w klatce piersiowej, duszność, parestezje), oceniane według CTCAE. Skuteczna była premedykacja i dalsze leczenie kontynuowano bez powikłań. Podkreśla to znaczenie CTCAE w dynamicznej ocenie objawów ostrych i adaptacji leczenia w czasie rzeczywistym.

Immunoterapia u dzieci z CHL (niwolumab + brentuksymab vedotin)[3]

W badaniu obejmującym dzieci z nawrotowym chłoniakiem Hodgkina stosowano CTCAE do oceny toksyczności nowego schematu leczenia immunoterapeutycznego. Pomimo skuteczności leczenia (100% całkowitych remisji), odnotowano jedno poważne działanie niepożądane. CTCAE okazała się narzędziem adekwatnym również w pediatrii, jednak dalsze badania mogą wymagać doprecyzowania progów toksyczności dla dzieci.

Ocena schematów limfodeplecji w terapii CAR-T[4]

W porównaniu dwóch schematów przygotowujących do terapii CAR-T (fludarabina + cyklofosfamid vs. bendamustyna), CTCAE v.5 posłużyła do oceny toksyczności hematologicznej, infekcji oraz zespołu CRS. Wskazano jednak, że bardziej zaawansowane skale (np. ASTCT) mogą być trudne do implementacji poza ośrodkami specjalistycznymi, co czyni CTCAE narzędziem bardziej uniwersalnym dla ośrodków mniej zaawansowanych technologicznie.

Zastosowanie CTCAE w badaniu pilotażowym z OMT[5]

W badaniu pilotażowym nad technikami osteopatycznymi u dzieci i młodych dorosłych z nowotworami, CTCAE użyto do oceny bezpieczeństwa terapii wspomagającej. Wyniki sugerują, że CTCAE jest użyteczna również w ocenie skutków ubocznych terapii nefarmakologicznych w onkologii.

W onkologii skale oceny toksyczności, takie jak RTOG-EORTC i LENT-SOMA, odgrywają istotną rolę w ocenie bezpieczeństwa i skutków ubocznych terapii, zwłaszcza w kontekście leczenia nowotworów układu nerwowego, chłoniaków oraz guzów klatki piersiowej. Poniżej krótkie notatki o analizowanych badaniach.

Skala RTOG-EORTC w leczeniu pierwotnego chłoniaka ośrodkowego układu nerwowego (PCNSL)[6]

W badaniu pilotażowym nad zastosowaniem temozolomidu (TMZ) w skojarzeniu z radioterapią (RT) u pacjentów z PCNSL, skala RTOG-EORTC została wykorzystana do oceny ostrej i późnej toksyczności leczenia. Badanie miało na celu określenie maksymalnej tolerowanej dawki TMZ, przy jednoczesnym zachowaniu skuteczności terapii. Większość pacjentów osiągnęła całkowitą lub częściową remisję, co potwierdza skuteczność zaproponowanego schematu. Jednocześnie wykazano przydatność skali RTOG-EORTC w kompleksowej ocenie toksyczności (wątrobowej, nerkowej, mózgowej, hematologicznej) – choć jej rutynowe zastosowanie może być ograniczone przez niedobory w dokumentacji klinicznej.

Skala LENT-SOMA w ocenie powikłań radioterapii

Skala LENT-SOMA została opracowana do oceny późnych efektów radioterapii i jest szeroko wykorzystywana w praktyce klinicznej. Znajduje zastosowanie w ocenie toksyczności narządowej.

- **Radioterapii klatki piersiowej[7]:**

Skala pozwala na monitorowanie kardiotoxyczności, stanowiącej istotne zagrożenie u pacjentów poddawanych napromienianiu tego obszaru. Umożliwia także porównywanie ryzyka sercowych powikłań w kontekście skojarzonego leczenia onkologicznego.

- **Oszczędzającej operacji piersi z radioterapią śródoperacyjną (IORT)[8]:**

U pacjentek z nawrotem raka piersi po wcześniejszym napromienianiu, zastosowanie skali LENT-SOMA pozwoliło na potwierdzenie bezpieczeństwa zabiegu oraz niskiego poziomu późnej toksyczności. Przy czym po 5 latach obserwacji odsetek pacjentek bez wznowy miejscowej wynosił 89,9%, a przeżycie ogólne – 82,7%, co pozwoliło ocenić w znacznym stopniu powikłania odległe.

- **Zmodyfikowanej terapii wysokodawkowej (HDT) w chłoniaku nieziarniczym (NHL)[9]:**

Skala była pomocna w ocenie toksyczności, szczególnie skórnej, w terapii łączącej całkowite napromienianie ciała (TBI) oraz napromienianie pola zaangażowanego (IF-RT). Choć leczenie charakteryzowało się wysoką skutecznością (przeżycie ogólne 100%, bez nawrotu 78%),

odnotowano także przypadki późnych powikłań (m.in. zespół mielodysplastyczny, zaburzenia hormonalne), co potwierdza konieczność długoterminowej obserwacji pacjentów.

Skala MAT (MASCC Antiemesis Tool)[10] była użyta do oceny nudności i wymiotów u pacjentów z nowotworami hematologicznymi poddawanych chemioterapii. Celem badania było zrozumienie częstości występowania tych objawów i ich zależności od rodzaju chemioterapii. Nudności były częstszym problemem niż wymioty, zwłaszcza w fazie opóźnionej po chemioterapii. Narzędzie to jest odpowiednie głównie dla **badania prospektywnych** – wymaga aktywnego raportowania objawów przez pacjentów, co jest trudne do odwzorowania w badaniach retrospektywnych opartych na dokumentacji medycznej. Ze względu na brak danych o samopoczuciu pacjenta w dokumentacji medycznej, skala MAT nie została uwzględniona jako narzędzie główne w ocenie toksyczności w omawianym badaniu nad HL.

Q-TWiST[11] to **ilościowa metoda oceny jakości życia**, która uwzględnia zarówno czas przeżycia, jak i okresy bez objawów choroby (TWiST), nawrotów (REL) oraz toksyczności leczenia (TOX). Umożliwia integrację **subiektywnej oceny jakości życia** z danymi klinicznymi, poprzez przypisanie wartości użyteczności (utility weights) poszczególnym stanom zdrowia. **Została** wykorzystana do oceny jakości życia pacjentów, w tym przypadku tych z chłoniakiem grudkowym (FL), leczeni schematami CVP i R-CVP. Celem analizy było porównanie jakości życia pacjentów w obu grupach, uwzględniając czas przeżycia, czas bez objawów choroby oraz toksyczności. Jednak ze względu na konieczność pozyskania odpowiednich współczynników jakości, oraz **szczególowych informacji o stanie pacjenta w czasie**, które często nie są odnotowywane w dokumentacji medycznej Q-TWiST jest mniej przydatna w badaniach retrospektywnych opartych na dokumentacji medycznej.

Faza 2 – Weryfikacja możliwych do oceny parametrów CTCAE

W drugiej fazie badania przeprowadzono analizę wyników laboratoryjnych i klinicznych uzyskanych od 20 pacjentów, mając na celu wyłonienie tych parametrów, które pojawiły się przynajmniej jednokrotnie u każdego z nich. Kryterium powtarzalności uznano za kluczowe, ponieważ jednorazowe, niestandardowe badania nie są wiarygodnym źródłem do oceny działań niepożądanych (AE – Adverse Events), a tym samym nie nadają się do ujęcia w algorytmie automatycznej detekcji.

Następnie, dla wyselekcjonowanych parametrów klinicznych przyporządkowano odpowiadające im jednostki klasyfikacyjne według skali CTCAE (Common Terminology Criteria for Adverse Events). Każdy z tych parametrów został poddany ocenie pod kątem zgodności z kryteriami prostoty – tj. możliwości jego oceny na podstawie rutynowo dostępnych wyników badań. Przykładowo, gorączka neutropeniczna wymaga jedynie danych o temperaturze ciała oraz poziomie neutrofili, które są standardowo monitorowane w codziennej praktyce klinicznej.

W odróżnieniu od parametrów bazujących na wartościach liczbowych, kryteria oparte na wynikach EKG były trudniejsze do oceny, ze względu na ich opisowy charakter oraz konieczność uwzględnienia informacji o interwencji (np. dla przypisania 2 lub 3 punktów). Mimo to, z uwagi na częste wykonywanie EKG u pacjentów z chłoniakiem Hodgkina (HL) i obecność opisów wyników w dokumentacji medycznej, zdecydowano o ich uwzględnieniu w analizie częstości występowania działań niepożądanych.

W przypadku wielu kategorii CTCAE przypisanie 5 punktów odpowiadało zgonowi, natomiast 4 punkty – stanom zagrożenia życia lub wymagającym pilnej interwencji. Uznano, że tego typu zdarzenia są na tyle poważne, iż ich wykrycie w dokumentacji jest łatwe i nie wymaga wysokiej czułości algorytmu na wczesnym etapie monitorowania.

Zdarzają się parametry w ctcae, gdzie ocena dla jednego z pięciu punktów jest łatwa dla regex, a innego nie. Dla przykładu, w eozynofilii przypisanie 3 punktów oznacza konieczność włączenia glikokortykosteroidów, co jest trudne do automatycznej oceny. Jednak poziom 1 punktu, definiowany jako przekroczenie wartości wyjściowej lub górnej granicy normy, jest możliwy do prostego oszacowania. Analogiczna logika doboru została zastosowana dla wszystkich parametrów – tak, aby przynajmniej jeden poziom punktowy był łatwy do oceny.

Wprowadzono również kilka wyjątków od tej zasady w przypadku parametrów trudnych do bezpośredniego uchwycenia. Przykładowo, hemoliza może być podejrzewana na podstawie wzrostu poziomu retikulocytów oraz bilirubiny, natomiast ostre uszkodzenie nerek wymaga danych z obserwacji klinicznych, choć możliwe jest zastosowanie wskaźników pośrednich (np. wzrost stężenia kreatyniny i mocznika) do wygenerowania alertu dla użytkownika.

Mimo że zespół lizy guza nie jest możliwy do jednoznacznej oceny na podstawie pojedynczego wyniku badania, często pojawia się we wzmiankach w historii choroby. Z tego względu uznano za zasadne przeanalizowanie również jego częstości występowania.

W przypadku EKG, mimo opisowego charakteru wyników, można łatwo zidentyfikować poważne zaburzenia rytmu, takie jak asystolia. Dodatkowo możliwa jest analiza semantyczna (słownikowa), umożliwiająca automatyczną klasyfikację danego opisu jako istotnego klinicznie (≥ 1 punkt) lub nieistotnego (0 punktów), co stanowi wartość dodaną przy automatycznej analizie danych.

Dla zawału mięśnia sercowego konieczne jest spełnienie dwóch warunków: zmiany w EKG oraz wzrost poziomu enzymów sercowych – oba te kryteria spełniają wymóg prostoty (maksymalnie dwa parametry wejściowe). Z kolei dla wydłużonego QT, konieczne jest jedynie wskazanie wartości QT w opisie EKG, co jest ocenialne. W przypadku niedociśnienia tętniczego, punktacja opiera się na wartościach ciśnienia krwi oraz obecności objawów i potrzeby interwencji. Nadciśnienie tętnicze natomiast wymaga zastosowania bardziej złożonych kryteriów, często zależnych od wieku pacjenta.

Podsumowując, selekcja parametrów do oceny w CTCAE została przeprowadzona tak, by zapewnić ich maksymalną zgodność z rutynowo dostępnymi danymi oraz z zachowaniem klinicznego znaczenia, przy jednoczesnym uwzględnieniu ograniczeń związanych z automatyzacją oceny, chcąc by jak najwięcej wyników oceniać za pomocą wyrażeń regularnych i algorytmów warunkowych, unikając konieczności analizy języka naturalnego.

Tabela 1. Wybrane parametry w skali CTCAE na podstawie analizy danych w dokumentacji medycznej

Nazwa badania	Powtórzenie	odpowiednik w ctcae	prostota oceny*	akceptacja**
Morfologia krwi (CBC)	tak	tak	tak	PRAWDA
mocznik	tak	tak	tak	PRAWDA
Kreatynina	tak	tak	tak	PRAWDA
Glukoza	tak	tak	tak	PRAWDA
AST	tak	tak	tak	PRAWDA
ALT	tak	tak	tak	PRAWDA
Kwas moczowy	tak	tak	tak	PRAWDA
Bilirubina całkowita	tak	tak	tak	PRAWDA
ALP (fosfataza zasadowa)	tak	tak	tak	PRAWDA
GGTP	tak	tak	tak	PRAWDA

Albuminy	tak	tak	tak	PRAWDA
Haptoglobina	tak	tak	tak	PRAWDA
Fibrynogen	tak	tak	tak	PRAWDA
RKZ (równowaga kwasowo-zasadowa)	tak	tak	tak	PRAWDA
APTT	tak	tak	tak	PRAWDA
INR	tak	tak	tak	PRAWDA
GFR (wskaźnik przesączania kłębuszkowego)	tak	tak	tak	PRAWDA
Badanie ogólne moczu	tak	tak	tak	PRAWDA
Triglicerydy	tak	tak	tak	PRAWDA
Mleczany	tak	tak	tak	PRAWDA
Kinaza kreatyninowa (CK)	tak	tak	tak	PRAWDA
Wapń skorygowany	tak	tak	tak	PRAWDA
Wapń jonowy	tak	tak	tak	PRAWDA
Masa ciała	tak	tak	tak	PRAWDA
Ciepłota ciała	tak	tak	tak	PRAWDA
Temperatura ciała	tak	tak	tak	PRAWDA
Na	tak	tak	tak	PRAWDA
K	tak	tak	tak	PRAWDA
Cl	tak	tak	tak	PRAWDA
Pi	tak	tak	tak	PRAWDA
Magnez (Mg^{2+})	tak	tak	tak	PRAWDA
EKG*	tak	tak	tak	PRAWDA
OB (Odczyn Biernackiego)	tak	nie	tak	FAŁSZ
CRP (Białko C-reaktywne)	tak	nie	tak	FAŁSZ
Wapń całkowity (Ca^{2+})	nie	tak	tak	FAŁSZ
HbA1c (hemoglobina glikowana)	nie	nie	tak	FAŁSZ
Prokalcytonina (PCT)	tak	nie	tak	FAŁSZ
RTG	tak	tak	nie	FAŁSZ
TK/CT	tak	tak	nie	FAŁSZ
MRI	nie	tak	nie	FAŁSZ
USG	tak	tak	nie	FAŁSZ
Badanie grupy krwi i przeciwciał	nie	tak	tak	FAŁSZ
Odleżyny	tak	nie	tak	FAŁSZ

Objaśnienie do tabeli

Tabela zawiera zestawienie badań laboratoryjnych i klinicznych analizowanych w kontekście ich przydatności do oceny działań niepożądanych według skali CTCAE.

- **Nazwa badania** – pełna nazwa parametru lub procedury medycznej.
- **Powtórzenie** – informacja, czy dane badanie było wykonywane wielokrotnie (co najmniej dwukrotnie) w przebiegu leczenia u pacjentów.
- **Odpowiednik w CTCAE** – określa, czy dany parametr ma przypisany bezpośredni odpowiednik w klasyfikacji działań niepożądanych CTCAE.
- ***Prostota oceny**** – ocena wykonalności automatycznej klasyfikacji danego badania według poniższych kryteriów:
 - badanie nie wymaga jednoczesnej analizy więcej niż dwóch różnych wyników,

- ocena nie wymaga pogłębionej wiedzy klinicznej, której nie da się uzyskać z epikryzy i dokumentacji postępowania medycznego,
- wynik badania ma charakter liczbowy lub jednoznaczny opisowo (np. EKG: „asystolia”).
- **Akceptacja** – końcowa kwalifikacja parametru do dalszej analizy. Parametr uznawano za zaakceptowany, jeśli spełniał wszystkie poniższe warunki:
 - był wykonywany co najmniej dwukrotnie w trakcie leczenia,
 - był możliwy do oceny według powyższych kryteriów prostoty,
 - posiadał odpowiednik w skali CTCAE.

Faza 3 – Projekt aplikacji webowej do oceny działań niepożądanych

W trzeciej fazie badania przeprowadzono pilotażową ocenę częstości oraz rodzaju występujących działań niepożądanych (DN) u pacjentów poddanych leczeniu przeciwnowotworowemu. Wśród najczęściej obserwowanych powikłań znalazły się: nadciśnienie tętnicze, przyrost masy ciała, spadek liczby limfocytów oraz neutrofilii, wzrost poziomu limfocytów, obniżenie poziomu leukocytów oraz hiperfosfatemia.

Dominującym schematem leczenia w analizowanej grupie był ABVD. Do oceny zakwalifikowano 5 pacjentów, w przedziale wiekowym 20–70 lat, obejmującym obie płcie.

Oceny DN dokonano z wykorzystaniem dedykowanego narzędzia webowego, dostępnego pod adresem: <https://git.1ioe.top/HL/components/ctcae.html>. Aplikacja działająca jako statyczna strona internetowa nie zapisująca żadnych danych w pamięci lokalnej urządzenia została opracowana w celu usprawnienia procesu identyfikacji i klasyfikacji działań niepożądanych na podstawie skali CTCAE, przy zapewnieniu jednocześnie bezpieczeństwa danym. Umożliwia ona użytkownikowi szybkie wyszukiwanie odpowiednich kategorii DN, generowanie gotowego raportu w formacie ułatwiającym transfer danych do arkusza kalkulacyjnego (np. Excel), a także zawiera formularze do wprowadzenia identyfikatora pacjenta oraz daty hospitalizacji – z możliwością zaszyfrowania i odszyfrowania przy użyciu hasła.

W badaniu zastosowano pełną anonimizację danych – żadne dane osobowe nie były wprowadzane do systemu. Strona aplikacji zawiera komunikat informujący użytkownika o zasadach ochrony danych osobowych, wyświetlany przy pierwszym uruchomieniu. W treści komunikatu użytkownik proszony jest o wyłączenie wszelkich niezaufanych wtyczek przeglądarki oraz o kodowanie danych wrażliwych, jeśli zostaną wprowadzone.

Najczęściej występujące DN obejmowały zaburzenia możliwe do oceny na podstawie podstawowych badań laboratoryjnych (morfologia krwi) oraz pomiaru masy ciała. Co istotne, wszystkie te działania niepożądane mogą być oceniane z wysoką dokładnością z wykorzystaniem prostego algorytmu analitycznego, bez konieczności angażowania modeli językowych (LLM).

Tabela 2. Najczęściej występujące działania niepożądane i ich klasyfikacja wg CTCAE

(Miejsce na tabelę z wynikami)

Faza 4 – Symulacja systemu AMMS

W ramach czwartej fazy opracowano statyczną stronę internetową, która w uproszczony sposób odwzorowuje mechanikę działania systemu AMMS (Advanced Medical Monitoring System). Aplikacja została zaimplementowana w języku polskim, co wynika z faktu, że dane wykorzystywane w fazach 2 i 3 również były pozyskiwane i analizowane w tym języku. Interfejs użytkownika oraz opisy przypadków pacjentów są zatem spójne językowo z pozostałymi etapami projektu.

Na potrzeby demonstracyjne przygotowano przykładowe dane trzech fikcyjnych pacjentów chorujących na chłoniaka Hodgkina. Dla każdego z pacjentów zaprezentowano informacje kliniczne, wyniki badań oraz przebieg leczenia. Strona umożliwia prześledzenie decyzji podejmowanych przez algorytm zaproponowany w tej publikacji (faza piąta) na podstawie wprowadzonych danych, co pozwala użytkownikowi lepiej zrozumieć jego logikę działania jako małoinwazyjne rozszerzenie do AMMS.

Symulacja nie posiada zaplecza serwerowego ani funkcji uczenia maszynowego – pełni jedynie rolę poglądowej wizualizacji. Jej głównym celem było przetestowanie koncepcji rozszerzenia interfejsu użytkownika oraz ocena intuicyjności obsługi przez osoby mające doświadczenie z podobnymi systemami medycznymi.

Dostęp do symulacji możliwy jest pod adresem:

<https://git.1ioe.top/HL/components/AMMS.html>

Faza 5 – Skrypt oceniający parametry CTCAE

W ramach piątej fazy projektu opracowano modułowy skrypt JavaScript służący do oceny parametrów CTCAE na podstawie danych tekstowych zawartych w elementach typu textarea w aplikacjach webowych, takich jak system AMMS. Skrypt został umieszczony w pojedynczym pliku .js, co znacząco upraszcza jego integrację – wymaga jedynie dodania poniższej liniiki w kodzie źródłowym aplikacji:

```
<script type="module" src="https://git.1ioe.top/HL/components/analyser.js"></script>
```

Dla celów demonstracyjnych rozwiązanie to zostało wdrożone na stronie internetowej autora.

W toku implementacji napotkano istotne wyzwania. Pierwszym z nich była ograniczona funkcjonalność na urządzeniach mobilnych, wynikająca z zastosowania skrótów klawiszowych jako podstawowego mechanizmu sterowania. Problem ten można rozwiązać poprzez modyfikację kodu skryptu przez dostawcę oprogramowania, w celu dostosowania go do środowiska mobilnego. Dla potrzeb symulacji, uznano, że wystarczające będzie użycie urządzeń stacjonarnych, także dlatego, że na takich działa system AMMS w większości szpitali.

Drugim wyzwaniem była niska czytelność algorytmów opartych na wyrażeniach regularnych, zarówno w części odpowiedzialnej za ekstrakcję danych wyjściowych (baselines), jak i w module oceniającym CTCAE. Utrudnia to analizę bezpieczeństwa oraz proces dostosowywania skryptu do konkretnej konfiguracji systemu medycznego. Z tego względu – ograniczono użycie RegExp – na korzyść mieszanego podejścia – proste regExp i JavaScript.

Skrypt obsługuje wszystkie parametry uznane za istotne w trzeciej fazie badania, klasyfikując je jako wymagające lub niewymagające danych bazowych (baselines). Po załadowaniu do dokumentu HTML, automatycznie tworzone są dwa panele: panel alertów oraz panel narzędziowy. Pierwszy z nich wyświetlany jest centralnie w oknie przeglądarki (50% szerokości i 50% wysokości viewportu), zaś drugi – dynamicznie w miejscu kliknięcia użytkownika w pole textarea.

Panel narzędziowy jest aktywowany przy każdym kliknięciu w edytowalny obszar tekstowy, zakładając, że to w nim znajdują się wyniki badań. W momencie aktywacji skrypt przeprowadza analizę tekstu – w przypadku wykrycia nieprawidłowości, wyświetlany jest również panel alertów. Oba panele wykluczają się nawzajem wizualnie – kliknięcie w jeden ukrywa drugi. Naciśnięcie przypisanego skrótu klawiszowego zamyka wszystkie aktywne panele.

Panel narzędzi zawiera trzy przyciski funkcyjne: (1) ikonę miotełki do czyszczenia pól komentarza z danych bazowych zapisanych przez skrypt, (2) ikonę edytora służącą do ekstrakcji danych bazowych z tekstu oraz (3) ikonę narzędzi umożliwiającą wstawienie danych w formacie JSON do komentarza na podstawie uprzednio pobranych informacji. W przypadku kliknięcia pola zawierającego już dane w formacie JSON, panel alertów nie zostanie wyświetlony, nawet jeśli pole zawiera treści mogące wskazywać na nieprawidłowe wyniki badań – co teoretycznie może prowadzić do problemów wskutek błędnego umieszczenia danych w polu z wynikami badań zamiast polu komentarza, choć w środowisku AMMS takie pola zazwyczaj mają status tylko do odczytu.

4. Dyskusja

Zautomatyzowana klasyfikacja zdarzeń niepożądanych przy użyciu skali CTCAE umożliwia szybkie i obiektywne wykrywanie klinicznie istotnych stanów wynikających z działań niepożądanych leczenia onkologicznego. Potwierdzają to wyniki badań [12-17], które wskazują na wysoką skuteczność narzędzi automatyzujących ten proces w porównaniu z oceną manualną. Potencjalne zastosowania takiego podejścia obejmują:

- wsparcie decyzji klinicznych,
- systemy ostrzegania w czasie rzeczywistym,
- monitorowanie pacjentów poza badaniami klinicznymi, zwłaszcza w sytuacjach wymagających intensywnego nadzoru.

Do ograniczeń badania należy zaliczyć m.in. zależność od kompletności danych oraz częsty brak integracji z pełnymi danymi klinicznymi, takimi jak wywiad lekarski, obserwacja kliniczna czy samoocena pacjenta.

Jak wcześniej wspomniano, celem niniejszego badania było zaprojektowanie małoinwazyjnego systemu, stanowiącego rozszerzenie istniejącego systemu AMMS. System ten miałby na bieżąco wyświetlać alerty podczas przeglądania dokumentacji medycznej. Aby zapewnić skuteczność działania, konieczne było zastosowanie powszechnie przyjętej skali – wybrano CTCAE – i opracowanie na jej podstawie algorytmu analizującego dane w czasie rzeczywistym. W tym celu skupiono się wyłącznie na tych obszarach skali CTCAE, które można skutecznie zintegrować z systemem o wspomnianych wcześniej cechach.

W literaturze nie znaleziono jednoznacznych kryteriów pozwalających na wybór najbardziej istotnych parametrów w CTCAE [13,14]. Aby uniknąć subiektywnego doboru, przeanalizowano manualnie, które obszary w tej skali najczęściej dają nieprawidłowości w wynikach badań pacjentów z chłoniakiem Hodgkina.

Parametry o najwyższej liczbie punktów odzwierciedlają te, które stanowią częste skutki uboczne standardowej terapii HL i jednocześnie mogą być ocenione na podstawie rutynowo zbieranych danych medycznych. W przypadku niektórych pozycji, szczególnie kardiologicznych, możliwa była jedynie częściowa

ocena – wyższe stopnie skali wymagają bowiem określenia ciężkości stanu pacjenta oraz oceny zagrożenia życia, co może być dokonane jedynie pośrednio, na podstawie dokumentacji.

Przykładowo, dla migotania przedsionków, CTCAE rozróżnia stopień 1 i 2 na podstawie obecności lub braku objawów. Dodatkowym wyzwaniem w hematologii jest fakt, że sama choroba może wpływać na wyniki morfologii krwi obwodowej, co utrudnia rozróżnienie pomiędzy działaniami niepożądanymi a objawami choroby podstawowej.

Warto także zauważyć, że niektóre nieprawidłowości w wynikach badań mogą się utrzymywać wskutek ich przeoczenia przez personel medyczny i braku wdrożenia leczenia objawowego – np. izolowana hiperfosfatemia. Choć ogranicza to możliwość wiarygodnej oceny częstości występowania działań niepożądanych w badaniach retrospektywnych, pozwala jednocześnie stwierdzić, które z nich miały charakter utrwalony, a nie przejściowy – złagodzony przez leczenie objawowe.

Podczas zbierania danych odnotowano także szereg problemów z oceną niektórych pozycji – jak opisano w części metodologicznej – takich jak trudności w interpretacji, konieczność dodatkowych badań czy niejednoznaczność danych. Umożliwiło to selekcję z pierwotnej listy tych parametrów, które są potencjalnie łatwe i szybkie do pełnej oceny, a tym samym nadają się do przekształcenia w algorytmy w języku JavaScript, zastosowane w niniejszym symulatorze.

Dzięki temu, wzorem innego badania [12], możliwe jest dokumentowanie działań niepożądanych w sposób wystandaryzowany, z minimalnym obciążeniem dla użytkownika, w oparciu o istniejącą dokumentację medyczną. Takie podejście sprzyja późniejszym analizom zbiorczym i metaanalizom.

W scenariuszu manualnej oceny działań niepożądanych istotnym problemem pozostaje duży nakład pracy oraz niska dokładność oceny efektów leczenia – w dużej mierze zależna od osoby oceniającej, jej doświadczenia oraz zmęczenia pracą. Potencjalnym rozwiązaniem tej trudności jest automatyczne raportowanie danych z dokumentacji medycznej. Przykładem takiego podejścia jest badanie [13], w którym oceniono zastosowanie tego typu rozwiązania w populacji pediatrycznej z ostrą białaczką szpikową (AML). Wykorzystano tam dane zanonimizowane w formacie CSV, z których za pomocą zapytań SQL wyodrębniono informacje niezbędne do uzupełnienia dziewięciu pól skali CTCAE.

Analizie poddano 49 pacjentów i wykazano, że zastosowanie algorytmicznej ekstrakcji danych poprawiło dokładność i precyzję oceny w skali CTCAE w porównaniu z oceną manualną, jednocześnie obniżając koszty procesu raportowania. Wyniki te są spójne z naszymi obserwacjami dotyczącymi oceny manualnej i automatycznej w ramach symulatora, gdzie zaobserwowano różnice w sposobach interpretacji wyników badań – zwłaszcza w rozróżnianiu nieprawidłowości wynikających z choroby podstawowej od tych będących skutkiem ubocznym leczenia.

Ciekawą obserwacją jest to, że dla oceny manualnej często stosowano heurystykę polegającą na sprawdzaniu jedynie wartości nieprawidłowych w wynikach badań, przy jednoczesnym uwzględnieniu progów ustalonych dla wystandaryzowanej populacji młodych dorosłych, co miało na celu skrócenie czasu oceny. Zarówno w naszym badaniu, jak i w omawianym badaniu [13], nie uwzględniono parametrów nienumerycznych – głównie ze względu na konieczność przetwarzania języka naturalnego oraz retrospektywny charakter analizy, w którym opisy kliniczne różnią się w zależności od osoby je sporządzającej.

W innym badaniu [14] wdrożono specjalnie dla tego celu zaprojektowany całościowy system raportowania działań niepożądanych. Choć jego zastosowanie zmniejszyło obciążenie pracą dzięki specyfice aplikacji, to

jednak nie w takim stopniu jak w badaniu pierwszym [13]. Warto zaznaczyć, że każdy system oceny działań niepożądanych – niezależnie od stopnia automatyzacji – zawiera punkty szczególnie wrażliwe na duplikację, brak kompletności lub błędne uzupełnienie danych wymaganych w skali CTCAE.

Źródłem tych problemów bywa m.in. niewłaściwa metoda oceny związku przyczynowo-skutkowego, rozbieżności terminologiczne pomiędzy CTCAE a dokumentacją medyczną, czy też różnice w funkcji personelu zgłaszającego AE – przykładowo pielęgniarki częściej raportują działania poważne, natomiast lekarze skupiają się na objawach łagodnych i przejściowych[14].

Możliwymi rozwiązaniami tego problemu są: (1) opracowanie systemu gromadzącego dane i oceniającego alerty w sposób zespołowy, poprzez udział wielu członków personelu – co jednak generuje znaczne koszty i nakład pracy, lub (2) zastosowanie algorytmów wspierających decyzje, opartych na wystandaryzowanych, obiektywnych badaniach diagnostycznych.

W omawianym kontekście wskazano również na istotny problem przetwarzania synonimów[14]. Choć wykorzystanie dużych modeli językowych (LLM) czyni ten proces stosunkowo prostym, to przy zastosowaniu metod tradycyjnych jest on znacznie bardziej skomplikowany. Proponowanym przez nas podejściem jest wykorzystanie LLM do stworzenia listy synonimów, które można następnie przetestować na zanonimizowanych danych, weryfikując ich skuteczność w identyfikowaniu odpowiednich pojęć.

Wydaje się to wykonalne, zważywszy na stosunkowo dobrze ustrukturyzowany, cyfrowy charakter danych przechowywanych w polskich systemach szpitalnych. W pełni zautomatyzowane przetwarzanie danych sprawdza się szczególnie dobrze w przypadku wyników laboratoryjnych – danych numerycznych. Natomiast dane wymagające oceny klinicznej (np. powikłania kardiologiczne) mogą być lepiej analizowane w modelu hybrydowym, opartym na interakcji z upoważnionym członkiem zespołu medycznego.

W przypadku badań laboratoryjnych, dodatkowe utrudnienia w ocenie wg CTCAE w wersji 5.0 stanowią tzw. *baselines* – czyli wartości wyjściowe, które należy porównywać z wynikami uzyskanymi przed rozpoczęciem leczenia. W jednym z badań [15] zastosowano w tym celu makro służące do analizy danych już przetworzonych do edytowalnej postaci w programie Excel. Idąc za tym przykładem, możliwe jest wstępne wyodrębnienie danych z dokumentacji medycznej, a następnie – w przetworzonym i lokalnie przechowywanym zbiorze – przeprowadzenie analizy za pomocą relatywnie prostych algorytmów.

W innej publikacji dokonano przeglądu automatycznych technik wykrywania i oceny działań niepożądanych (AE). Wśród najważniejszych wniosków wskazano duże zróżnicowanie w częstości rejestrowanych AE oraz niedostateczne opisy stosowanych metod, co istotnie utrudniało ich porównanie [16].

Obecnie nadal dominuje podejście oparte na ręcznym zgłaszaniu działań niepożądanych, jednak dopiero implementacja metod automatycznego wykrywania [17] umożliwiła znaczące zwiększenie wykrywalności AE. Narzędzia tego typu – oparte na wyzwalaniu regułowym – uznaje się obecnie za złoty standard. Warto jednak zwrócić uwagę na jedną z kluczowych kwestii podniesionych w badaniu [16], a mianowicie dużą liczbę fałszywych alarmów. Można im potencjalnie zapobiec poprzez zastosowanie progów istotności dla detekcji. Drugim istotnym problemem okazała się nie zawsze kompletna dokumentacja medyczna – co można częściowo poprawić, integrując dane z różnych źródeł.

Alternatywne podejście zaprezentowano w pracy [17], gdzie wykorzystano słowniki terminologii medycznej oraz dwukierunkowy klasyfikator typu BERT, służący do wykrywania wzmianek o działaniach niepożądanych. W tym modelu maszyna osiągnęła 80% trafności, podczas gdy człowiek – 90%. Istnieje wysokie prawdopodobieństwo, że wykorzystanie dużych modeli językowych (LLM) pozwoliłoby na dalsze

zwiększenie skuteczności tego typu analiz. Należy jednak pamiętać, że wiązałoby się to również z większym ryzykiem ujawnienia danych wrażliwych podmiotom zewnętrznym.

Systematyczna dokumentacja toksyczności leczenia przeciwnowotworowego w warunkach ambulatoryjnych stanowi jeszcze większe wyzwanie niż w warunkach stałego pobytu w szpitalu, zwłaszcza gdy nie towarzyszy jej dedykowana infrastruktura organizacyjna i techniczna. W literaturze podejmowano próby uproszczenia tego procesu, między innymi poprzez wprowadzenie skali PRO-CTCAE – kwestionariusza samooceny wypełnianego bezpośrednio przez pacjentów. Przykładowo, zastosowanie formularzy online zawierających 32 pytania pozwalało na bieżące monitorowanie dolegliwości związanych z leczeniem [12]. Niemniej jednak, wyniki tego typu badań wskazują na ograniczoną skuteczność – w jednej z analiz, pomimo zgłoszeń toksyczności, nie odnotowano żadnych zmian w sposobie leczenia (0%). Co więcej, aż 50% pacjentów uznało, że obowiązek każdorazowego wypełniania kwestionariusza podczas wizyt ambulatoryjnych jest dla nich nieakceptowalny [12].

Efektywne wdrożenie systemu opartego na PRO-CTCAE wymaga zapewnienia pacjentom możliwości wygodnego dostępu do formularzy – czy to poprzez ich własne urządzenia mobilne, czy udostępnione komputery w placówkach medycznych. Mimo to, nawet przy spełnieniu tych warunków, pozostaje trudność związana z oceną toksyczności u pacjentów o ograniczonej zdolności do komunikacji – spowodowanej np. zaburzeniami funkcji poznawczych lub barierą językową. Dodatkową przeszkodą jest fakt, że skala PRO-CTCAE powstała w języku angielskim, a jej adaptacja do innych języków wymaga zasobów finansowych oraz zaangażowania specjalistów w zakresie tłumaczeń medycznych.

W tym kontekście, wartościowym podejściem komplementarnym może być wykorzystanie klasycznej skali CTCAE w połączeniu z automatyczną analizą wyników badań laboratoryjnych oraz zapisów dokumentacji medycznej. Taka metoda pozwala na ocenę skutków ubocznych leczenia w sposób bardziej obiektywny i niezależny od aktywnego udziału pacjenta, stanowiąc istotne wsparcie dla zespołów klinicznych w przypadkach, gdzie samoocena nie jest możliwa lub nie przynosi wiarygodnych rezultatów.

5. Wnioski

Zaproponowane w niniejszej pracy podejście umożliwia wdrożenie półautomatycznego systemu wykrywania działań niepożądanych (AE) w oparciu o skalę CTCAE, bez konieczności głębokiej integracji z istniejącymi systemami informatycznymi szpitali (takimi jak HIS czy AMMS) oraz bez istotnego zwiększenia obciążenia pracą personelu medycznego. Kluczowym elementem tego rozwiązania jest możliwość przeprowadzania analizy danych lokalnie – po stronie klienta – co w sposób znaczący podnosi poziom bezpieczeństwa informacji medycznych i minimalizuje ryzyko ich nieuprawnionego udostępnienia.

Zastosowanie prostych narzędzi przetwarzania tekstu, takich jak wyrażenia regularne (*regex*), okazuje się być efektywną alternatywą dla kosztownych i zasobożernych modeli językowych LLM (Large Language Models), szczególnie w jednostkach ochrony zdrowia dysponujących ograniczoną infrastrukturą technologiczną. Dzięki temu możliwe jest uniknięcie problemu nadmiernego „monitorowania” parametrów klinicznie nieistotnych, które mogą być błędnie interpretowane przez modele LLM. Co więcej, ograniczenie się do lekkich technologii eliminuje konieczność utrzymywania złożonej i kosztownej infrastruktury obliczeniowej – zarówno lokalnej, jak i chmurowej – przy jednoczesnym zachowaniu wysokiej skuteczności wykrywania AE.

W celu przezwyciężenia problemu ograniczonej liczby manualnie przeanalizowanych danych pacjentów, co negatywnie wpływa na jakość systemów detekcji, warto uwzględnić metodę *Ripple Down Rules* (RDR). Umożliwia ona iteracyjne rozszerzanie algorytmu o nowe reguły, na podstawie interakcji z użytkownikiem, który może wskazać nieprawidłowości lub nowe wzorce w danych. Dzięki temu system staje się coraz bardziej precyzyjny i dopasowany do lokalnych warunków oraz typowych przypadków klinicznych.

Obecnie zaprezentowany algorytm ma charakter półautomatyczny, co pozwala na jego szybkie wdrożenie bez modyfikacji infrastruktury szpitalnej. Niemniej jednak, możliwe jest jego pełne zintegrowanie z systemami typu AMMS, co zwiększyłoby poziom automatyzacji oraz efektywność całego procesu. Istnieje również możliwość przepisania algorytmu do innych języków programowania oraz jego integracja z pozostałymi modułami systemów zarządzania danymi pacjentów. Takie podejście może znacząco przyczynić się do poprawy jakości monitorowania bezpieczeństwa terapii w placówkach medycznych, bez ponoszenia nadmiernych kosztów technologicznych i organizacyjnych.

6. Bibliografia

1. <https://redbook.streamliners.co.nz/5802.pdf>
2. Comer H, Cardwell K. Brentuximab Vedotin Infusion Reaction Management: A Case Study. *J Adv Pract Oncol*. 2017 Sep-Oct;8(6):626-629. Epub 2017 Sep 1. PMID: 30310723; PMCID: PMC6167082.
3. Greve P, Beishuizen A, Hagleitner M, Loeffen J, Veening M, Boes M, Peperzak V, Diez C, Meyer-Wentrup F. Nivolumab plus Brentuximab vedotin +/- bendamustine combination therapy: a safe and effective treatment in pediatric recurrent and refractory classical Hodgkin lymphoma. *Front Immunol*. 2023 Jul 31;14:1229558. doi: 10.3389/fimmu.2023.1229558. PMID: 37583696; PMCID: PMC10423930.
4. Ghilardi G, Chong EA, Svoboda J, Wohlfarth P, Nasta SD, Williamson S, Landsburg JD, Gerson JN, Barta SK, Pajarillo R, Myers J, Chen AI, Schachter L, Yelton R, Ballard HJ, Hodges Dwinal A, Gier S, Victoriano D, Weber E, Napier E, Garfall A, Porter DL, Jäger U, Maziarz RT, Ruella M, Schuster SJ. Bendamustine is safe and effective for lymphodepletion before tisagenlecleucel in patients with refractory or relapsed large B-cell lymphomas. *Ann Oncol*. 2022 Sep;33(9):916-928. doi: 10.1016/j.annonc.2022.05.521. Epub 2022 Jun 9. PMID: 35690221; PMCID: PMC11332249.
5. Belsky JA, Brown AM. Investigating the safety and feasibility of osteopathic manipulative medicine in hospitalized children and adolescent young adults with cancer. *J Osteopath Med*. 2024 Apr 29;124(9):399-406. doi: 10.1515/jom-2024-0013. PMID: 38669608.
6. Chiesa, S., Hohaus, S., Falcinelli, L. *et al.* Chemoradiotherapy with temozolomide after high-dose methotrexate for primary CNS lymphoma: a multicenter phase I study of a response-adapted strategy. *Ann Hematol* **99**, 2367–2375 (2020). <https://doi.org/10.1007/s00277-020-04220-1>
7. Mantini G, Smaniotto D, Balducci M, Dinapoli N, Campitelli M, Corvari B, Simili A, Ciarniello V. Radiation-induced cardiovascular disease: impact of dose and volume. *Rays*. 2005 Apr-Jun;30(2):157-68. PMID: 16294909.
8. Thangarajah F, Heilmann J, Malter W, Kunze S, Marnitz S, Mallmann P, Wenz F, Sperk E. Breast conserving surgery in combination with intraoperative radiotherapy after previous external beam

therapy: an option to avoid mastectomy. *Breast Cancer Res Treat.* 2018 Apr;168(3):739-744. doi: 10.1007/s10549-017-4657-y. Epub 2018 Jan 15. PMID: 29335922.

9. Ott M, Schmidberger H, Wörmann B, Albrecht CF, Pradier O, Hess CF. Involved-field irradiation in combination with total-body irradiation (TBI) does not increase short-term toxicity compared to TBI alone in patients with advanced-stage low-grade non-Hodgkin lymphoma. *Strahlenther Onkol.* 2002 May;178(5):245-51. doi: 10.1007/s00066-002-0932-9. PMID: 12082683.
10. Oiwa K, Hosono N, Itoh K, Ookura M, Matsuda Y, Tai K, Ueda T, Yamauchi T. [Evaluation of Chemotherapy-Induced Nausea and Vomiting in Patients with Hematological Malignancies Using MASCC Antiemesis Tool(MAT)]. *Gan To Kagaku Ryoho.* 2018 Jan;45(1):45-50. Japanese. PMID: 29362306.
11. Marcus R, Aultman R, Jost F. A quality-adjusted survival analysis (Q-TWiST) of rituximab plus CVP vs CVP alone in first-line treatment of advanced follicular non-Hodgkin's lymphoma. *Br J Cancer.* 2010 Jan 5;102(1):19-22. doi: 10.1038/sj.bjc.6605443. Epub 2009 Nov 17. PMID: 19920818; PMCID: PMC2813752.
12. Wang, Tian & Samuel, Joseph & Brown, M. & Vennettilli, Ashlee & Solomon, Hannah & Eng, Lawson & Liang, Mindy & Gill, Gursharan & Merali, Zahra & Tian, Chenchen & Cheng, Nicholas & Campbell, Matthew & Patel, Devalben & Liu, Ai & Liu, Geoffrey & Howell, Doris. (2018). Routine Surveillance of Chemotherapy Toxicities in Cancer Patients Using the Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). *Oncology and Therapy.* 6. 10.1007/s40487-018-0065-7.
13. Miller TP, Li Y, Getz KD, Dudley J, Burrows E, Pennington J, Ibrahimova A, Fisher BT, Bagatell R, Seif AE, Grundmeier R, Aplenc R. Using electronic medical record data to report laboratory adverse events. *Br J Haematol.* 2017 Apr;177(2):283-286. doi: 10.1111/bjh.14538. Epub 2017 Feb 1. PMID: 28146330; PMCID: PMC5384859.
14. Lencioni A, Hutchins L, Annis S, Chen W, Ermisoglu E, Feng Z, Mack K, Simpson K, Lane C, Topaloglu U. An adverse event capture and management system for cancer studies. *BMC Bioinformatics.* 2015;16 Suppl 13(Suppl 13):S6. doi: 10.1186/1471-2105-16-S13-S6. Epub 2015 Sep 25. PMID: 26424052; PMCID: PMC4597098.
15. Widel, Mario. "Implementing laboratory toxicity grading for CTCAE Version 5."
16. Musy S, Ausserhofer D, Schwendimann R, Rothen H, Jeitziner M, Rutjes A, Simon Trigger Tool–Based Automated Adverse Event Detection in Electronic Health Records: Systematic Review *Med Internet Res* 2018;20(5):e198 URL: <https://www.imir.org/2018/5/e198> DOI: 10.2196/jmir.9901
17. Classen DC, Resar R, Griffin F, Federico F, Frankel T, Kimmel N, Whittington JC, Frankel A, Seger A, James BC. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff (Millwood).* 2011 Apr;30(4):581-9. doi: 10.1377/hlthaff.2011.0190. Erratum in: *Health Aff (Millwood).* 2011 Jun;30(6):1217. PMID: 21471476.