

Prof. zw. dr hab. med. Marek Sasiadek  
Kierownik Katedry Radiologii  
i Zakładu Radiologii Ogólnej, Zabiegowej i Neuroradiologii  
Uniwersytetu Medycznego im. Piastów Śląskich we Wrocławiu  
ul. Borowska 213  
50-556 Wrocław

Wrocław, 5.10.2020

## RECENZJA

**lek. Charlie Alexander HAMM**

**rozprawa doktorska**

**“Development of an Interpretable Liver Tumor Diagnosis Tool using Deep Learning”**

Przedstawiam recenzję rozprawy doktorskiej zgodnie z Umową o Dzieło nr 64/2020 z Pomorskim Uniwersytetem Medycznym w Szczecinie.

1) Przedstawienie informacji o ocenianej rozprawie doktorskiej.

a) Tytuł rozprawy doktorskiej stanowiącej podstawę ubiegania się w aktualnym postępowaniu o nadanie stopnia doktora:

**“Development of an Interpretable Liver Tumor Diagnosis Tool using Deep Learning”**

b) Ocena układu rozprawy doktorskiej, w tym informacje o poszczególnych częściach składowych:

Rozprawa ma następujące części składowe:

- strona tytułowa

- informacja o przygotowaniu rozprawy w oparciu o tematycznie spójny zestaw publikacji:

**Hamm Charlie A**, Clinton J Wang, Lynn J Savic, Marc Ferrant, Isabel Schubert, Todd Schlachter, MingDe Lin et al: Deep learning for liver tumor diagnosis part I: development of a convolutional neural network classifier for multi-phasic MRI. *European Radiology* 2019; 29 (7): 3338-3347 (IF: 3,962; MNiSW 25 pkt lista “A”)

Wang Clinton J, **Charlie A Hamm**, Lynn J Savic, Marc Ferrant, Isabel Schubert, Todd Schlachter, MingDe Lin et al: Deep learning for liver tumor diagnosis part II: convolutional neural network interpretation using radiologic imaging features. *European Radiology* 2019; 29 (7): 3348-3357 (IF: 3,962; MNiSW 25 pkt lista “A”).

W obu artykułach było dwóch równoważnych pierwszych autorów: **Charlie A Hamm** oraz **Clinton J Wang**. **Niewymienieni w tej informacji współautorzy** to: James S. Duncan, Jeffrey C. Weinreb, Julius Chapiro i Brian Letzen.

Do wersji drukowanej rozprawy nie dołączono oświadczenia Doktoranta o udziale w przygotowywaniu w/wym. artykułów, ani oświadczeń drugiego równoważnego pierwszego autora i pozostałych autorów na temat udziału lek. Charlie A Hamma w przygotowaniu w/wym. artykułów. Oświadczenia te, na moją prośbę skierowaną do Przewodniczącego Rady Dyscypliny Nauki Medyczne Pomorskiego Uniwersytetu Medycznego, zostały dosłane przez wszystkich współautorów. **Spowodowało to opóźnienie mojej recenzji.**

Z przesłanych Oświadczeń, których kopie dołączam do recenzji, wynika, że Doktorant wniósł znaczący wkład na wszystkich etapach przygotowania publikacji; ponadto wszyscy współautorzy wyrazili zgodę na wykorzystanie obu artykułów w rozprawie doktorskiej.

- Spis treści (**bez numerów stron poszczególnych części; numerów stron nie ma zresztą w całej rozprawie**)

- Lista skrótów
- Wstęp (niecałe 2 strony)
- Cel badania (5 linijek)
- Materiał i metody (nieco ponad 5 stron), z 6 podrozdziałami i 2 tabelami
- Wyniki(około 4 strony) z 3 podrozdziałami, 2 tabelami i 3 rycinami
- Dyskusja (nieco ponad 5 stron)
- Wnioski (8 linijek tekstu)
- Streszczenie angielskie (ok. 1,5 strony)
- Streszczenie polskie (ok. 1,5 strony)
- Piśmiennictwo (50 pozycji)
- Wydruk dwóch artykułów wymienionych wyżej
- Podziękowania, m.in. dla: promotora rozprawy doktorskiej; Clintona Wanga (równoważny pierwszy autor w/wym. artykułów) oraz niektórych innych współautorów tych artykułów.

Układ rozprawy jest poprawny. Moje zastrzeżenia formalne to: 1. brak umieszczenia wszystkich współautorów artykułów cyklu doktorskiego w informacji na początku rozprawy (zwłaszcza, że jeden z niewymienionych był autorem korespondencyjnym obu artykułów!); 2. brak numeracji stron (zob. wyżej).

c. Ocena zastosowanego piśmiennictwa w ramach rozprawy doktorskiej

Piśmiennictwo składa się z 50 pozycji, w tym licznych z ostatnich kilku lat; wiele pozycji pochodzi z w/wym. dwóch artykułów, same te artykuły również są na liście piśmiennictwa rozprawy doktorskiej. Wszystkie 50 pozycji są cytowane w tekście rozprawy.

Ocena: zastosowane piśmiennictwo nie budzi zastrzeżeń.

d. Wskazanie oraz ocena celu pracy kandydata

Cel pracy został określony przez Doktoranta jako ocena algorytmu sztucznej inteligencji w postaci interpretowalnego systemu głębokiego uczenia (*deep learning*) w diagnostyce guzów wątroby w rezonansie magnetycznym (MR) w porównaniu z tradycyjną oceną radiologiczną, a także jego roli w procesie podejmowania decyzji co do postępowania lekarskiego.

Cel ten należy ocenić jako interesujący, gdyż rozwijanie algorytmów sztucznej inteligencji stanowi aktualnie jeden z najistotniejszych tematów w radiologii. Szkoda, że cele nie zostały przedstawione w dwóch punktach, tak jak w streszczeniu rozprawy; co byłoby bardziej przejrzyste.

e. Wskazanie oraz ocena zastosowanych metod badawczych

Rozprawa oparta jest na materiale 494 zmian wątroby (u 296 pacjentów) wykrytych w badaniu MR jamy brzusznej, podzielonych na 6 grup: naczyniak jamisty, ogniskowy przerost guzkowy (NFH), torbiel, wewnątrzwątrobowy rak dróg żółciowych (ICC), przerzut raka jelita grubego, rak wątrobowo-komórkowy (HCC). Obrazy tych zmian zostały wprowadzone do stworzonej w tym celu konwolucyjnej sieci neuronowej (*convolutional neural network- CNN*) i poddane analizie komputerowej metodą głębokiego uczenia (*deep learning- DL*), przy czym większość z nich (434) przeznaczono do zestawu treningowego (*training set*), a pozostałe (po 10 losowo wybranych zmian z każdej podgrupy) do zestawu testowego (*test set*).

Doktorant i współpracownicy w artykułach cyklu doktorskiego bardzo szczegółowo przedstawiają poszczególne etapy procesu analizy obrazów metodą *deep learning*, poczynając od protokołu akwizycji badania MR, poprzez metody postprocessingu uzyskanych obrazów, tworzenie modelu *deep learning* i przeprowadzenie uczenia w tym modelu. Następnie oceniono czułość i swoistość głębokiego uczenia w podziale na grupę treningową (*training set*) i testową (*test set*) i porównano je z wynikami oceny dwóch doświadczonych radiologów.

Generalnie metody badawcze są dobrze przemyślane i zaplanowane, oparte na bardzo dobrym przygotowaniu autorów, zwłaszcza w zakresie technik komputerowych. Natomiast słabszą stroną, o której zresztą pisze Doktorant, jest brak potwierdzenia histopatologicznego większości analizowanych zmian, co spowodowało, że ostateczne rozpoznanie było oparte

głównie na typowych cechach obrazowych i klinicznych. Drugim mankamentem jest fakt, że badania MR były wykonywane na różnych aparatach (o natężeniu pola magnetycznego 1,5 i 3 Tesle), co budzi wątpliwość o jednorodność analizy komputerowej.

Tym niemniej uważam, że część metodyczna rozprawy została przygotowana starannie i jest podstawą do uzyskania wiarygodnych wyników.

f. Ocena części rozprawy doktorskiej dotyczące omówienia wyników badań

Doktorant przedstawia w podrozdziale „Dyskusja” omówienie wyników badań, które jest uzupełnieniem omówienia (*Discussion*) w obu artykułach cyklu.

Doktorant zwraca uwagę, że klasyfikacja zmian w wątrobie za pomocą stworzonego przez autorów artykułów cyklu modelu DL (*deep learning*) pozwoliła na uzyskanie lepszych parametrów czułości, swoistości i dokładności diagnostycznej niż w przypadku oceny przez 2 doświadczonych radiologów. I tak dokładność w różnicowaniu 6 grup patologii wyniosła odpowiednio 90% vs 80/85%; określenie w klasyfikacji LI-RADS odpowiednio 92% vs 88%. System DL wykazał szczególnie wyższe niż radiolodzy wskaźniki czułości w rozpoznawaniu HCC, co wg Doktoranta jest związane z sugerowaniem się przez radiologów bardziej lub mniej jednoznacznym obrazem klinicznym.

W „Omówieniu” lek. Hamm analizuje też liczbę cech obrazowych, a w przyszłości także klinicznych która powinna być wprowadzona do analizy przez system DL. Zwraca uwagę, że z jednej strony może to zwiększyć efektywność systemu, lecz z drugiej strony zbyt duża liczba cech wprowadzonych do systemu może doprowadzić do obniżenia dokładności diagnostycznej.

W opinii Doktoranta system sztucznej inteligencji może potencjalnie stać się pomocniczym narzędziem dla radiologa jako „druga opinia”. Szansą dla zwiększenia akceptacji lekarzy dla takiego rozwiązania, jest proponowany przez autorów cyklu system oceny punktowej każdej poszczególnej cechy obrazowej.

W końcowej części „Omówienia” lek. Hamm zwraca uwagę na ograniczenia badania przeprowadzonego przez Niego i współautorów, m. in. stosunkowo małą liczbę zmian w poszczególnych podgrupach, konieczność wybrania jedynie przypadków o typowych cechach obrazowych, czy małą liczbę pacjentów z marskością wątroby. Dlatego postuluje dalsze badania nad rozwinięciem modelu CNN/DL w ocenie zmian w wątrobie, ze starannym dobraniem cech obrazowych i klinicznych, który będzie analizowany i trenowany przez algorytmy komputerowe.

Ogólnie rozdział „Omówienie” („*Discussion*”) rozprawy doktorskiej oceniam pozytywnie. Autor w sposób zwarty, lecz wyczerpujący (bardziej dokładnie niż w artykułach cyklu)

analizuje uzyskane wyniki pod względem zarówno możliwości samego systemu, jak jego zastosowania klinicznego. Uczciwie zwraca uwagę na ograniczenia badań będących podstawą rozprawy i proponuje kierunek dalszych badań, mających na celu doskonalenie systemu.

h. Informacja o ewentualnych nieprawidłowościach, które pojawiły się ocenianej rozprawie doktorskiej.

Po dostaniu Oświadczeń Doktoranta i współautorów artykułów cyklu doktorskiego, w których potwierdzony został samodzielny wkład lek. Hamma w artykułach cyklu, rozprawa spełnia warunki formalne. Nie stwierdzam więc nieprawidłowości formalnych, jednak mam kilka uwag krytycznych:

- zwyczajowo cykl doktorski liczy nie mniej niż 3 publikacje; w tym przypadku są to 2 artykuły, a de facto jeden artykuł podzielony na 2 części, opublikowane w tym samym czasopiśmie jedna po drugiej. Pod tym względem mimo pozytywnej oceny merytorycznej (potwierdzonej akceptacją artykułów przez renomowane czasopismo), mam poczucie pewnego dyskomfortu,

- moje zdziwienie budzi, że Doktorant dopiero po mojej interwencji przesłał oświadczenia współautorów artykułów cyklu doktorskiego. Jako naukowiec powinien od razu poprosić współautorów o oświadczenia, gdyż ich brak mógłby potencjalnie skutkować zastrzeżeniami prawnymi współautorów,

- za nieeleganckie i nieprofesjonalne uważam niewymienienie wszystkich współautorów artykułów cyklu doktorskiego (w tym autora korespondencyjnego) w informacji na początku rozprawy,

- streszczenie rozprawy nie jest w pełni spójne z tekstem rozprawy i artykułów cyklu, gdyż nie są w nim uwzględnione wszystkie istotne elementy rozprawy. Dodatkowo w streszczeniu polskim niektóre określenia tłumaczone z angielskiego są niezręczne, np. „sieć neuronów konwulsyjnych” zamiast prawidłowego określenia „konwolucyjna sieć neuronowa”

- drobnym, ale istotnym uchybieniem jest wspomniany już brak numeracji stron, co utrudnia zapoznanie się z rozprawą i sprawia, że spis treści staje się mało użyteczny.

i. Ocena, czy rozprawa doktorska stanowi oryginalne rozwiązanie problemu naukowego

W mojej opinii, pod względem merytorycznym tekst rozprawy, a zwłaszcza przedstawia oryginalny i istotny problem naukowy, który jest przeprowadzony za pomocą poprawnej metodyki i właściwej analizy wyników.

j. Ocena, czy rozprawa doktorska prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinie albo dyscyplinach oraz umiejętność samodzielnego prowadzenia pracy naukowej lub artystycznej

Uważam, że Doktorant wykazał się dobrą wiedzą w zakresie radiologii klinicznej jamy brzusznej oraz podstaw sztucznej inteligencji jako metody wspomaganie procesu diagnostycznego: zarówno jak współautor projektu opublikowanego w artykułach cyklu doktorskiego, jak autor pozostałej części tekstu rozprawy.

2) W przypadku, gdy rozprawę doktorską stanowi samodzielna i wyodrębniona część pracy zbiorowej, recenzja powinna zawierać ocenę indywidualnego wkładu kandydata w powstanie tej pracy.

Wkład Kandydata został określony w „Oświadczeniach” Doktoranta i współautorów artykułów cyklu doktorskiego. Zgodnie z tymi oświadczeniami wkład lek. Hamma w powstanie artykułów cyklu polegał na:

- przygotowaniu aplikacji do komisji bioetycznej,
- koncepcji badań na podstawie piśmiennictwa,
- określeniu aspektów klinicznych ogniskowych zmian wątroby pod kątem hipotezy badawczej,
- wyboru metod badawczych (m. in. kryteriów włączenia i wykluczenia, kryteriów analizy obrazu i parametrów dokładności zastosowanych metod),
- selekcji badań obrazowych w systemie PACS,
- postprocessingu badań obrazowych,
- uczestniczeniu w przygotowaniu konwolucyjnej sieci neuronalnej (wspólnie z równorzędnym pierwszym autorem publikacji, Clintonem Wangiem),
- przeprowadzenie procesu oceny radiologicznej badań MR,
- analizy i interpretacji uzyskanych wyników pod kątem ich znaczenia klinicznego,
- przygotowaniu rycin i tabel,
- napisania i rewizji manuskryptów,
- koordynacji procesu publikacji artykułów w „European Radiology”.

Wszyscy współautorzy wyrazili zgodę na wykorzystanie obu artykułów w rozprawie doktorskiej lek. Charlie Hamma.

Na podstawie oświadczeń Doktoranta i współautorów artykułów cyklu doktorskiego, można stwierdzić, że lek. Charlie Hamm miał decydujący udział w przeprowadzeniu badań i ich publikacji w artykułach cyklu doktorskiego, a tym samym przygotowana na ich podstawie rozprawa doktorska jest samodzielnym dokonaniem Doktoranta.

3) W przypadku gdy autor recenzji stwierdzi, że rozprawa doktorska zasługuje na wyróżnienie, należy wskazać uzasadnienie postawionego wniosku,

Pomimo mojej ogólnie wysokiej ocenie merytorycznej rozprawy, ze względu na wskazane powyżej uwagi krytyczne, nie wnioskuję o wyróżnienie rozprawy doktorskiej

Inne uwagi:

We „Wnioskach” Doktorant stwierdza, że dokładność interpretowalnego systemu głębokiego uczenia w ocenie zmian ogniskowych wątroby przewyższa wskaźniki uzyskane przez doświadczonych radiologów, a tym samym może być potencjalnie zastosowany w procesie decyzyjnym postępowania z pacjentem; wymaga to jednak dalszych badań prospektywnych.


Wnioski są sformułowane poprawnie i stanowią bezpośrednią odpowiedź na zaplanowane Cele rozprawy doktorskiej.

Podsumowując uważam, że rozprawa doktorska lek. Charlie Hamma jest wartościowym, rzetelnie zrealizowanym opracowaniem nowatorskiego zagadnienia, przynoszącego wyniki ważne pod względem poznawczym i klinicznym. Tworzące ją dwa artykuły opublikowane w renomowanym czasopiśmie radiologicznym stanowią zwartą całość, opartą o starannie przemyślane hipotezy badawcze i cele, zrealizowaną rzetelnie pod względem metodycznym. Doktorant miał wiodący udział w tych publikacjach, co potwierdzają oświadczenia współautorów. Kilka uwag krytycznych, które zawarłem w tekście recenzji, nie zmienia mojej całościowej wysokiej oceny pracy.

**Stwierdzam, że przedstawiona mi do oceny rozprawa doktorska lek. Charlie Hamma pt. “Development of an Interpretable Liver Tumor Diagnosis Tool using Deep Learning” spełnia warunki określone w art. 18 Ustawy z dnia 20 lipca 2018- Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2020 r., poz. 85 z późn. zm.).**

W związku z tym zwracam się do Wysokiej Rady Naukowej Dyscypliny Nauki Medyczne Pomorskiego Uniwersytetu Medycznego w Szczecinie o dopuszczenie lek. Charlie Hamma do dalszych etapów przewodu doktorskiego.

Prof. dr hab. med. Marek Sęsiadek

  
Prof. dr hab. med. Marek Sęsiadek  
Specjalista radiolog  
52-414 Wrocław, ul. Śniegockiego 33