

2019 -08- 23 *W.B. ul.*

prof. dr hab. Paweł Olko
Instytut Fizyki Jądrowej PAN
Radzikowskiego 152
31-342 Kraków

Recenzja rozprawy doktorskiej pani mgr Marty Giżyńskiej zatytułowanej

„Analysis of the influence of patient setup errors on the doses received by them during radiotherapy”

Zasadniczy postęp w radioterapii, jaki dokonał się w bieżącym stuleciu, jest wynikiem wprowadzanie nowych modalności promieniowania jak i bardziej precyzyjnych technik ich użycia. Postulat podania zadanej dawki do objętości leczonej i jej minimalizacji dla tkanek zdrowych jest możliwy do realizacji tylko w przypadku precyzyjnie przeprowadzonego ułożenia pacjenta. Jest to szczególnie istotne w przypadku narządów ruchomych, których kształt i położenie może ulec zmianie. Dlatego w radioterapii stosuje się techniki weryfikacji ułożenia pacjenta, mające zminimalizować związane z tym niepewności. Procedury te mogą być czasochłonne i niejednokrotnie wiązać się z wyższym obciążaniem radiacyjnym pacjenta np. na skutek częstego stosowania stożkowej tomografii komputerowej (CBCT). Dlatego podejmowane są badania, mające na celu optymalizację metod i częstości ułożenia pacjenta.

Celem pracy doktorskiej pani mgr Marty Giżyńskiej było ulepszenie protokołu weryfikacji ułożenia pacjenta oraz zaproponowanie nowej metody wyznaczania marginesu terapeutycznego, uwzględniającego międzyfrakcyjne trendy czasowe. Taka indywidualizacja ułożenia i planowania leczenia miała prowadzić do zmniejszenia niepewności podania dawek. Praca została wykonana pod kierunkiem pana prof. Pawła Kukołowicza z Instytutu Marii Skłodowskiej Curie- Centrum Onkologii w Warszawie (COI) i pana dr hab. Macieja Kamińskiego z Wydziału Fizyki Uniwersytetu Warszawskiego (UW) w jednym z czołowych ośrodków uprawiających fizyki medyczną w Polsce. Autorka pracy, która 13 lat temu zakończyła studia magisterskie, jest doświadczonym fizykiem medycznym.

Punktem wyjścia pracy była analiza dwóch baz danych pacjentów po radioterapii raka prostaty, zawierających zmierzone odchylenia od wartości referencyjnych w trzech wymiarach podczas podania kompletu frakcji. 100 pacjentów Centrum Onkologii w Warszawie było układanych na struktury kostne a 835 pacjentów ośrodka Erasmus RC w Rotterdamie na złote znaczniki. Na wstępie Autorka analizowała, czy rozkłady średnich niepewności ułożenia pacjentów $\overline{m}_p$ i ich rozkłady standardowe SD_p wykazują rozkłady normalne w całej populacji pacjentów oraz jakie są ich trendy czasowy podczas kolejnych frakcji. Dla populacji pacjentów z bazy COI hipoteza zerowa, że błędy systematyczne posiadają rozkład normalny nie mogła zostać odrzucona w przedziale ufności 99%. Jednakże, rozkład odchyłeń standardowych SD_p dla populacji pacjentów nie może być traktowany jako rozkład normalny dla około 40% przypadków. W przypadku analizy trendów czasowych około 50% przypadków rozkład nachylenia prostej regresji numer frakcji nie jest rozkładem normalnym. Podobna analiza przeprowadzona dla zależności czasowych bazy Erasmus MC pokazała, że ułożenia na złotych znacznikach prowadziła do mniejszej częstości pacjentów (ok. 20%) dla których rozkłady niepewności ułożenia nie były normalne. Ponieważ wiele niezależnych czynników wpływa na błąd ułożenia należy spodziewać się, że rozkład błędów ułożenia będzie rozkładem normalnym. Rozkłady

Pdk

odchylenia standardowego SD_p dla obu baz danych również nie okazały się normalne. Podobnie brak rozkładu normalnego wykazano dla rozkładów $\overline{m_p}$ dla trzech kierunków bazy Erasmus i dwóch COI. Dlatego, dla dopasowania rozkładów $\overline{m_p}$ zastosowano uogólniony rozkład logistyczny natomiast dla rozkładu odchyłeń standardowych rozkład Burra czyli uogólniony rozkład log-logistyczny.

Częstym zabiegiem wykorzystywanym w pracy była analiza statystyczna na próbie z rozkładu syntetycznie przygotowanej populacji, zbudowanej na podstawie parametrów rozkładów z obu baz danych. Aby rozróżnić trendy czasowe wynikające ze statystyki i z fizjologii pacjenta porównano rozkłady współczynników kierunkowych prostych utworzonych dla syntetycznej populacji bez trendu czasowego z rozkładami uzyskanymi klinicznie. Wstępna analiza obu baz danych bazy pokazała, że dla ponad połowy populacji błędy ułożenia zwiększały się dla kolejnych frakcji. Dlatego w dalszych częściach pracy podjęto analizę trendów czasowych błędów ułożenia.

W rozdziale 4 zastosowano dwa różne podejścia do parametryzacji rozkładów błędów ułożenia – klasyczne, w którym do parametryzacji stosowano rozkład normalny lub rozkład Burra, oraz poprzez parametryzację przy zastosowaniu linii trendu, gdzie średni błąd ułożenia jest liniową funkcją numeru frakcji. Rozkłady średniego błędu ułożenia uzyskane wprost z bazy danych Erasmus oraz przy zastosowaniu analizy trendy czasowego pozostawały w znakomitej zgodności. Aby zminimalizować błędy ułożenia stosuje się w onkologii protokoły weryfikacyjne, polegające na weryfikacji on-line (sprawdzanie pozycji przed każdym napromienianiem) lub off-line, w której decyzje podejmuje się na podstawie uprzednio zgromadzonej statystyki. Ilościowa ocena jakości parametryzacji została przeprowadzona z zastosowaniem protokołów weryfikacyjnych NAL (No Action Level) i eNAL (extended No Action Level). Protokół NAL zaleca korekcję błędu systematycznego po trzech pierwszych sesjach terapeutycznych natomiast w protokole eNAL błędy są dodatkowo korygowane pod koniec kolejnych tygodni terapii. Oba protokoły zastosowano zarówno do analizy populacji syntetycznych jak i danych klinicznych z baz Erasmus i COI. Wykazano, że klasyczna parametryzacja prowadzi do zaniżenia o 50-70% wariancji błędów systematycznych Σ_{res} dla populacji co prowadziłoby do zastosowania zbyt małych marginesów przy napromienianiu. Gdy odchylenie standardowe średniego błędu ułożenia pacjenta SD_p rośnie, maleje jakość przybliżenia Σ_{res} . W syntetycznej populacji bez uwzględnienia czynnika czasu zakłada się stałość SD_p . Ciekawym zabiegiem było zbadanie, jak odwrócenie kolejności frakcji wpływa na Σ_{res} . Dla populacji syntetycznej nie uległo ono zmianie natomiast dla klinicznej bazy Erasmus inwersja kolejności danych prowadziła do obniżenia Σ_{res} , szczególnie dla niewielkiej liczby frakcji. Autorka zauważyła, że w takim przypadku przy analizie procesu napromienienia nie powinno się zakładać stałości odchylenia standardowego rozkładu wartości średniej.

Błędy ułożenia wpływają na rozkłady dawek w objętości tarczowej oraz w narządach krytycznych. W metodzie van Herka marginesy dla objętości CTV-PTV wyznacza się nie indywidualnie na podstawie wyników pomiarów dla danego pacjenta ale na podstawie wyników uzyskanych dla populacji pacjentów. Szereg proponowanych metod optymalizacji procesu planowania leczenia wymaga wbudowanie ich do systemów TPS. W rozdziale 5 Autorka podjęła próbę indywidualizacji marginesów dla poszczególnych pacjentów. Punktem wyjścia była próba przewidzenia sumarycznego rozkładu dawki w PTV oraz narządach krytycznych dla różnych technik terapeutycznych, takich jak 3D-CRT, IMRT czy VMAT.

Zastosowano tu dwie metody obliczeń. W pierwszej, sumaryczny rozkład dawki wyliczano po wyliczeniu średniego błędu ułożenia po n frakcjach. Dawka ta była wyliczana dla pozycji przesuniętej o wektor m_{pf} . Inna metoda polegała na uśrednieniu rozkładu dawek w każdej z początkowych n frakcji. Ilościową ocenę tych dwóch metod badano poprzez ocenę zgodności sumarycznego rozkładu dawki z planem leczenia przy użyciu testu gamma, przez porównanie odpowiednich histogramów dawka objętość (DVH) (porównując D_{min} , D_{mean} i $D_{98\%}$) oraz przy zastosowaniu testów statystycznych. Wyniki testów gamma dla obu metod dawały bardzo zbliżone wyniki, z tym, że dla najbardziej konformalnej techniki VMAT potrzebna była najmniejsza liczba frakcji do uzyskania indeksu gamma 2 mm 2% mniejszego od 1 dla przynajmniej 95% wokseli. Wynik testu Manna-Whitneya pokazał, że dla liczby frakcji mniejszej niż $n=10$ nie można odrzucić hipotezy zerowej o braku różnicy wyznaczania sumarycznego rozkładu dawki obydwoma metodami. Interesującą konkluzją tej analizy jest stwierdzenie, że VMAT jest techniką najbardziej odporną na błędy ułożenia jeśli chodzi o geometryczny rozkład dawki. Obliczenia były przeprowadzone dla grupy 25 pacjentów COI dla których doświadczeni planiści przygotowali retrospektywnie plany dla technik 3D-CRT, IMRT i VMAT. Ograniczeniem powyższej analizy był brak danych dotyczących zmiennej anatomii prostaty, które potencjalnie można uzyskać np. z dziennej analizy skanów tomografii stożkowej CBCT. Autorka przekonuje, że zaproponowana w rozdziale 5 metoda obliczania skumulowanego rozkładu dawki może zostać zastosowana w algorytmach planowania leczenia dostosowanych dla indywidulanego pacjenta. Zaletą zaproponowanej metody jest możliwość jej zastosowania „na zewnątrz” systemu planowania leczenia a więc niezależnie od oprogramowania, którym dysponuje planujący.

Jedną z istotnych osiągnięć rozprawy jest zaproponowanie modyfikacji protokołu weryfikacji protokołu ułożenia pacjenta No Action Level, NAL. Zasadniczym pomysłem Autorki jest indywidualizacja protokołu NAL poprzez identyfikację pacjentów, dla których ryzyko niepoprawnej estymacji średniego błędu ułożenia jest największe i określanych jako „large movers”. W nowym protokole, określonym jako individualized No Action Level (iNAL) korekta ułożenia jest obliczana po pierwszych 3 frakcjach. Po każdej kolejnej wyznacza się odchylenie standardowe średniej z wszystkich poprzednich frakcji $SEM_{mp}(n)$ i porównuje z arbitralnie wybraną wartością punktu odcięcia, którego przekroczenie nakazuje wykonanie kolejnej weryfikacji ułożenia. Wybór punktu odcięcia zaproponowano badając krzywe ROC, przedstawiające zmienność wychwycenia klasy faktycznie pozytywnej (TPR) w zależności od poziomu błędu popełnianego na klasie faktycznie negatywnej (TNR). Pokazano, że protokół iNAL prowadzi do niższej, w porównaniu z protokołem NAL, zmienności błędu systematycznego Σ_{res} przy takiej samej liczbie zdjęć weryfikacyjnych. W pracy zamieszczono również przepis zastosowania protokołu iNAL.

W rozdziale 7 pracy zaproponowano nowe podejście do określenia marginesów terapeutycznych dla objętości CTV-PTV, uwzględniające czasowe trendy międzyfrakcyjne. Błędy ułożenia podzielono na dwie grupy: podczas wykonywania napromieniania (błędy przypadkowe) oraz podczas przygotowań (błędy systematyczne). Trend czasowy jest traktowany jako błąd systematyczny wokół którego istnieje rozkład błędów przypadkowych. Dla wyznaczenia składowej marginesu dla błędu systematycznego zdefiniowano wielkość nazwaną maksymalnym odchyleniem (Maxium Deviation, MD), na które składa się średni błąd ułożenia pacjenta oraz komponent zależny od nachylenia linii trendu oraz liczby frakcji. Dla każdej osi margines jest wyliczany niezależnie. Zaletą metody, w porównaniu do stosowanej

od kilkunastu lat metody van Herka jest uwzględnienie trendu czasowego, który wpływa na wielkość błędów przypadkowych. Weryfikacja metody została przeprowadzona poprzez wyznaczenie frakcji pacjentów dla których błąd systematyczny ułożenia nie mieścił się w wyliczonym marginesie błędu. Weryfikacje przeprowadzono zarówno dla danych klinicznych z bazy Erasmus jak i dla utworzonej populacji syntetycznej. Stosowanie metody ułatwia fakt, że przygotowano stabelaryzowane wartości marginesów dla różnych wartości błędu systematycznego oraz dla różnych trendów czasowych. Znacznie tej części pracy polega m.in. na tym, że Autorka pokazała, że dotychczas stosowana metoda van Herka daje zbyt małe marginesy w przypadku wystąpienia trendu czasowego dla kolejnych frakcji napromieniania.

Praca pisana jest z dużą starannością należną tekstom matematycznym, za co Autorce należą się słowa uznania. Jednak tekst jest dość oszczędny co przy obszernej notacji wymaga od czytającego szczególnej uwagi. Jako przykład podam rozważania w rozdziale 7, w którym definiuje trend czasowy błędu jako $x_{p,i}$ dla danego pacjenta p i kierunku i , gdy chwilę potem oblicza margines prawdopodobieństwa błędu ułożenia dla kierunku x , które zależy od wartości x_i^2 , x_j^2 i x_k^2 . Zajmuje trochę czasu ustalenie, że x i $x_{p,i}$ oraz $x_{p,i}$ i x_i^2 nie mają ze sobą wiele wspólnego.

Praca jest bardzo starannie przygotowana pod względem edycyjnym i językowym. Nieliczne przypadki, które zdaniem recenzenta są potknięciami językowymi, wyliczone są poniżej:

Str.13, l. 7 od końca Gross Tumor Volume (...) consist -> consists

Str. 25, l.4 od końca Ditribution -> Distribution

Str. 73, l.14 o dołu: Presented results does not -> do not

Str.97, l.2 od dołu "... other n first treatment fractions" -> during n first

Str. 105, l.1 "...it would be very important" -> would be very important

Str.113, l.9 od dołu "...has be in upper" -> has to be in upper

Str. 136, l.1 The proposed novel recipe is based of a similar procedure -> on a similar

Str. 136 l. 9 "...within each all systematic setup errors" -> within all systematic setup

Str. 136 l. 0 from CT -> from CTV

Dogłębna analiza statystyczna wyników off-line ułożenia pacjentów jest ważnym zagadnieniem fizyki medycznej, szczególnie dla nowotworów prostaty, dla których stosuje się dość powtarzalne schematy planowania i ułożenia pacjenta. Współczesne trendy w radioterapii idą jednak w kierunku weryfikacji ułożenia pacjenta przy każdym polu terapeutycznym czy wręcz w trakcie napromieniania (IGRT, 4DCT), tak jak ma to np. miejsce w czasie napromieniania na Cyber Knifie. Na razie rozdzielczość przestrzenna tomografii stożkowej (CBCT) nie dorównuje konwencjonalnej tomografii komputerowej ale rozwój technologiczny jest tak szybki, że metody weryfikacji off-line będą tracić swe dotychczasowe znaczenie. Autorka jest tego świadoma i w paru miejscach pracy daje temu wyraz, m.in. pod koniec dyskusji rozdziału 7.

Te wszystkie uwagi nie wpływają w najmniejszym stopniu na moją bardzo wysoką ocenę przedstawionej rozprawy. Przedstawia ona autentycznie nowe wyniki w dziedzinie fizyki medycznej, mogące znaleźć zastosowanie w praktyce klinicznej. Stwierdzam, że przedstawiona mi do recenzji rozprawa doktorska pani mgr Marty Giżyńskiej spełnia z nadmiarem wszelkie ustawowe i zwyczajowe wymagania stawiane tego typu pracom. Dlatego też wnoszę do Rady

PWW

Wydziału Fizyki UW o dopuszczenie pani mgr Marty Giżyńskiej do dalszych etapów przewodu doktorskiego.

Autentycznymi osiągnięciami pracy jest zaproponowanie nowego protokołu weryfikacji *off-line*, nazwanego individualized No Action Level Protocol (iNAL) oraz przedstawienie przepisu na obliczanie marginesów terapeutycznych uwzględniających międzyfrakcyjne trendy czasowe. Ta druga metoda została zgłoszona do publikacji w renomowanym czasopiśmie „Medical Physics”. W związku z tym stawiam wniosek o wyróżnienie rozprawy.



Paweł Olko

Kraków, 18 sierpnia 2019 roku