

## Porównanie algorytmów detekcji fali V w badaniu słuchowych potencjałów wywołanych pnia mózgu

### Comparison of wave V detection algorithms in auditory brainstem responses

**Bartosz Trzaskowski**

Instytut Fizjologii i Patologii Słuchu, Światowe Centrum Słuchu, Zakład Audiologii Eksperymentalnej, Warszawa/Kajetany

**Adres autora:** Bartosz Trzaskowski, Światowe Centrum Słuchu, Zakład Audiologii Eksperymentalnej, ul. Mokra 17, Kajetany, 05-830 Nadarzyn, e-mail: b.trzaskowski@gmail.com

#### Streszczenie

Słuchowe potencjały wywołane pnia mózgu (ang. *Auditory Brainstem Responses*, ABR) są zapisami zmian potencjału elektrycznego rejestrowanego na powierzchni głowy w odpowiedzi na bodziec słuchowy. Odzwierciedlają synchroniczną aktywność neuronów nerwu VIII i kolejnych etapów drogi słuchowej w pniu mózgu. Obecnie ABR stanowią jedną z najczęściej stosowanych metod badania słuchu. Znajdują zastosowanie w estymacji progu słyszenia, monitoringu funkcji nerwu słuchowego i pnia mózgu podczas operacji, różnicowaniu zaburzeń słuchu oraz badaniach przesiewowych. W pracy przedstawione i porównane zostały wybrane systemy automatycznej detekcji ABR, opracowane i opisane w czasopismach naukowych przez różne zespoły badawcze od czasu odkrycia słuchowych potencjałów wywołanych pnia mózgu.

**Słowa kluczowe:** słuchowe potencjały wywołane pnia mózgu • automatyczna detekcja odpowiedzi

#### Abstract

Auditory brainstem responses (ABRs) are changes in the electric potential recorded on scalp caused by the auditory stimulus. They reflect synchronous activity of neurons in VIII nerve and subsequent segments of auditory pathway in brainstem. Nowadays ABRs are one of the most common methods of objective hearing testing. They are useful in objective threshold estimation, monitoring of function of auditory nerve and brainstem during surgery, differentiation of hearing disorders and auditory screening. In this paper, selected systems of automatic ABR detection described in scientific journals by different research teams since the discovery of the auditory brainstem responses were presented and compared.

**Key words:** auditory brainstem responses • automatic response detection

#### Wstęp

Słuchowe potencjały wywołane pnia mózgu (ang. *Auditory Brainstem Responses*, ABR) zostały po raz pierwszy opisane w 1970 r. przez Dona Jewetta, Michaela Romano i Johna Willistona [1,2]. Badacze rejestrowali pierwszych 10 ms odpowiedzi od momentu podania bodźca słuchowego typu trzask i zasugerowali, że zarejestrowana odpowiedź pochodzi z pnia mózgu. W pracy przedstawiono odpowiedzi uzyskane dla sześciu osób ze słuchem normalnym, wraz z oznaczeniami głównych fal.

Obecnie ABR stanowi jedną z najczęściej stosowanych obiektywnych, czyli niewymagających współpracy ze strony osoby badanej, metod badania słuchu [3]. Znajdują zastosowanie między innymi w estymacji progu słyszenia, monitoringu funkcji nerwu słuchowego i pnia mózgu

podczas operacji, różnicowaniu zaburzeń słuchu czy badaniach przesiewowych. W przypadku większości standardowo wykonywanych badań ABR wynik jest ustalany na podstawie wzrokowej analizy zapisu przez osobę wykonującą badanie. Poprawna ocena wyników badań ABR nie jest prosta i wymaga doświadczenia lub przynajmniej gruntownego przeszkolenia.

Trudność w ocenie wyników ABR bierze się przede wszystkim z tego, że potencjały wywołane pnia mózgu są odpowiedziami bardzo słabymi. Ich amplituda jest rzędu dziesiątych części  $\mu\text{V}$ , co powoduje, że są podatne na różnego rodzaju zakłócenia mające swoje źródło w otoczeniu lub organizmie osoby badanej. Zakłócenia pochodzące z sieci energetycznej, związane z falami elektromagnetycznymi w otoczeniu, lub artefakty związane z aktywnością mięśniową nakładają się na zapis ABR, utrudniając ocenę wyniku.

Bardzo poważnym zakłóceniem w badaniach ABR jest spontaniczna elektryczna aktywność mózgu (ang. *Electroencephalography*, EEG), której nie da się uniknąć podczas rejestracji. Jest ona bardzo silna w porównaniu z odpowiedzią ABR i może charakteryzować się amplitudą do 100  $\mu$ V. W skali czasowej charakterystycznej dla ABR, czyli do ok. 20 ms, spontaniczna aktywność mózgu ma charakter losowego szumu.

Wkład takich niechcianych rodzajów aktywności i zakłóceń może zostać zmniejszony poprzez zastosowanie procesu uśredniania wielu pojedynczych rejestracji lub odrzucanie pomiarów z artefaktami. Jednak nie wszystkie artefakty da się w ten sposób usunąć. Zakłócenia związane z bodźcem, takie jak np. fala sonomotoryczna, zostaną przez uśrednianie dodatkowo wzmocnione [4].

Mnogość czynników zakłócających pomiary ABR powoduje olbrzymią różnorodność morfologii odpowiedzi i dlatego może ona być trudna do rozpoznania, w szczególności w przypadku zapisów dla okołoprogowych natężeń bodźca słuchowego. Dokładność wyniku może zależeć od kwalifikacji osoby oceniającej badanie lub jej aktualnej kondycji psychofizycznej.

W 2004 r. Michael Vidler i David Parker [5] podjęli próbę określenia wpływu wzrokowej oceny wyniku badania na zgodność rezultatów. W eksperymencie udział wzięło 16 ekspertów z wieloletnim doświadczeniem w audiologii (średnio 15 lat) i przeprowadzaniu oraz ocenie badań ABR (ponad 8 lat). Uczestnikom przedstawiono 12 szeregów natężeniowych wytwarzanych w komputerowej symulacji badania. Ich zadaniem polegało na określeniu progów odpowiedzi. Wyniki tego eksperymentu pokazały, że wzrokowa ocena wyników badania ABR może mieć duży wpływ na ostateczny wynik. Nie było ani jednego szeregu natężeniowego, dla którego eksperci byłiby zgodni w ocenie progów. Największa różnica pomiędzy ocenami ekspertów wynosiła 60 dB, a w przypadku 9 z 12 szeregów rozbieżność sięgała przynajmniej 35 dB.

Opracowanie systemu automatycznej detekcji odpowiedzi ABR pozwoliłoby zmniejszyć ryzyko występowania różnic w ocenie wyników. W chwili obecnej różne ośrodki mogą kierować się odmiennymi preferencjami przy ocenie wyników. System automatyczny mógłby stanowić podstawę do standaryzacji kryteriów oceny. Stworzenie systemu wspomagającego osobę oceniającą zapis ABR wpłynęłoby na skrócenie czasu oceny wyniku, co bezpośrednio przełożyłoby się na zmniejszenie kosztów badania.

Próby opracowania systemu automatycznej detekcji ABR podejmowane są od ponad 30 lat. Przez ten czas wiele zespołów badawczych stosowało wiele różnych podejść. Testowane były metody oparte na wzorcach odpowiedzi, kryteriach, statystyce, analizie sygnałów czy sieciach neuronowych. Niniejsza praca motywowana jest chęcią zebrania w jednym miejscu i porównania wybranych prób opracowania systemu automatycznej detekcji ABR.

### Algorytmy detekcji fali V

W 1979 r. Claus Elberling do detekcji ABR zastosował wzorce odpowiedzi [6]. Zapisy ABR w grupie 24 pacjentów

z czysto ślimakowymi ubytkami słuchu zostały dostosowane w amplitudzie i latencji i wykorzystane do uzyskania normatywnych zapisów ABR służących jako wzorce odpowiedzi. Algorytm wykorzystuje miarę podobieństwa zapisów i wzorców. Dla zarejestrowanych odpowiedzi ABR obliczana była funkcja korelacji z wzorcem dla odpowiedniego natężenia bodźca. Obliczano dwa parametry: maksymalną wartość korelacji  $R$  i przesunięcie tej wartości w latencji  $\Delta L$  i na podstawie tych dwóch wielkości wnioskowano o obecności odpowiedzi i jej latencji. Autorzy wyznaczyli na płaszczyźnie  $[R, \Delta L]$  elipsy, wewnątrz których mieściło się 95% przypadków zapisów, co pozwalało z istotnością statystyczną 5% stwierdzić, czy badany zapis zawiera szukaną odpowiedź. Według autorów kliniczne zastosowanie metody dało zachęcające rezultaty, jednak w pracy brak jest danych liczbowych na temat efektywności systemu. Taka metoda rozpoznawania wzorców wydaje się dobra w przypadku szukania odpowiedzi określonej postaci, jednak zapisy ABR spotykane w typowej pracy klinicznej charakteryzują się różną morfologią.

W innej pracy z 1980 r. Weber i Fletcher zaproponowali system niewrażliwy na dużą różnorodność odpowiedzi ABR [7], oparty na obliczeniach współczynnika korelacji pomiędzy dwoma zapisami. Metoda bazowała na założeniu powtarzalności poszczególnych składowych i stopniowym spadku podobieństwa zapisów wraz ze zmniejszaniem natężenia bodźca i zbliżaniem się do progów ABR. Rejestrowano zapis 10 ms (256 punktów czasowych) od momentu podania bodźca, a współczynnik korelacji obliczano w przedziale 5–10 ms, gdzie odpowiedź jest najsilniejsza i gdzie powinna znajdować się fala V. W celu zwiększenia skuteczności metody oprócz dwóch odpowiedzi na bodziec rejestrowane były też dwa zapisy kontrolne bez podania bodźca. Zamiast prostego wnioskowania na podstawie korelacji pomiędzy dwiema odpowiedziami obliczano współczynnik korelacji pomiędzy sygnałami, ale dodatkowo współczynnik korelacji obu sygnałów z sygnałami kontrolnymi. Następnie wykonywana była transformacja  $Z'$ -Fischera i statystyczne oszacowanie, czy otrzymana korelacja pomiędzy zapisami z odpowiedzią jest znacząco wyższa od korelacji z sygnałami kontrolnymi bez odpowiedzi. Działanie algorytmu przetestowano na grupie 12 osób dorosłych o słuchu w normie. Otrzymywane średnie wartości współczynnika korelacji wynosiły od  $R=0,86$  dla 60 dB SL do  $R=-0,05$  dla  $-10$  dB SL. System uzyskał jeszcze 83% wykryć odpowiedzi dla niskiego natężenia 15 dB SL, a także nie wygenerował wyników fałszywie dodatnich.

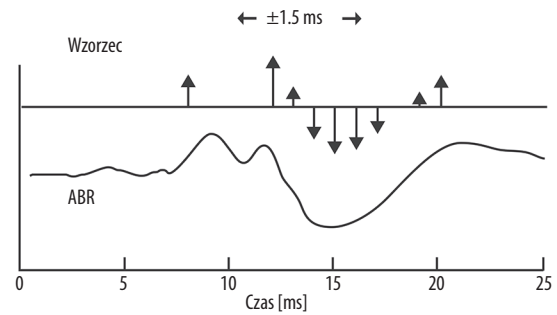
Do roku 1984 badania ABR przeprowadzano, wykonując określoną liczbę pojedynczych rejestracji lub przerywając test, gdy zdaniem osoby wykonującej badanie odpowiedź była już wystarczająco widoczna. Ponieważ siła odpowiedzi i poziom szumu zależą od bardzo wielu czynników, takich jak na przykład grubość czaszki pacjenta, impedancja elektrod czy wielkość zakłóceń pochodzących z otoczenia lub z organizmu samego pacjenta, to taki sposób rejestracji ABR nie był w stanie zapewnić odpowiedniej określonej minimalnej jakości pomiaru. W 1984 r. Don i Elberling [8,9] zaproponowali miarę nazwaną przez nich  $F_{sp}$ , będącą stosunkiem wariancji po czasie uśrednionego zapisu i wariancji w pojedynczym punkcie czasowym po pojedynczych rejestracjach. Z matematycznego punktu widzenia nie ma znaczenia punkt czasowy, w którym jest

obliczane  $F_{sp}$ , natomiast istotne jest, aby znajdował się w oknie, w którym zachodzi uśrednianie po czasie. Zastosowanie  $F_{sp}$  pozwala w trakcie pomiaru oszacować stosunek energii sygnału do szumu (ang. *Signal-to-noise ratio*, SNR) i wykonać taką liczbę uśrednień, która zapewni określoną minimalną jakość odpowiedzi. Prace te nie prezentują konkretnego algorytmu wykrywania odpowiedzi, jednak w kontekście obiektywizacji i standaryzacji metod pomiaru ABR są to bardzo ważne publikacje. Miara  $F_{sp}$  jest do dziś standardowo stosowana w wielu urządzeniach klinicznych.

Inny sposób określenia stosunku energii sygnału do szumu, oparty na sumie i różnicy dwóch jednocześnie rejestrowanych sygnałów [10,11], wykorzystany w systemie automatycznej detekcji ABR Mason [12–14]. System wymagał dwóch buforów zawierających 19,2-milisekundowe zapisy, będące średnią z 2048 lub 4096 pojedynczych odpowiedzi. Rejestracja wykonywana była równocześnie – naprzemiennie parzyste pojedyncze odpowiedzi sumowane były w pierwszym buforze, nieparzyste w drugim. Lokalizacja fali V wykrywana była jako pozycja maksymalnego podobieństwa obu zapisów poprzez obliczenie współczynnika korelacji w oknie długości 9 ms przesuwany równoległe w obu buforach. Stosunek sygnału do szumu estymowano na podstawie pomiaru amplitudy odpowiedzi w określonym punkcie największego podobieństwa i amplitudy szumu określonej jako różnica zapisów w obydwu buforach. System mógł określić wynik badania na trzy sposoby: odpowiedź obecna, odpowiedź prawdopodobnie obecna i brak odpowiedzi, w zależności od tego, czy wcześniej założone wartości współczynnika korelacji, stosunku amplitud odpowiedzi do szumu i wielkości amplitudy odpowiedzi zostały osiągnięte. W trybie przesiewowym decyzja, czy wynik przesiewu jest dodatni czy ujemny, podejmowana była na podstawie dwóch z tych trzech kryteriów. Dodatkowym mechanizmem kontroli warunków pomiaru i wiarygodności oceny zapisu było oszacowanie amplitudy szumu i odpowiedzi. W przypadku niespełnienia wymaganych kryteriów badanie było powtarzane.

Przebadano grupę 13 dorosłych osób ze słuchem normalnym (7 mężczyzn i 6 kobiet) w wieku od 17 do 42 lat (średni wiek – 27 lat) oraz 30 dzieci i niemowląt (19 chłopców i 11 dziewczynek) z podejrzeniem różnych stopni ubytków słuchu, w wieku od dwóch miesięcy do 16 lat (średni wiek – 2,5 roku). Osoby dorosłe były badane w pozycji siedzącej w dźwiękoszczelnym pomieszczeniu. Niemowlęta i dzieci były badane, kiedy siedziały, leżały na kanapie, były trzymane przez rodziców lub siedziały u nich na kolanach. Sedacja zastosowana była tylko w 4 przypadkach, gdy dziecko było bardzo aktywne.

Dokładność systemu w oznaczaniu progu została porównana z oznaczeniami operatora dla 25 badań progowych w grupie dorosłych o słuchu normalnym i 50 badań w grupie dzieci i niemowląt z podejrzeniami ubytku słuchu. W grupie dorosłych w 92% przypadków różnice w oznaczeniach progu przez system i przez operatorów były nie większe niż 10 dB, a w grupie dzieci odsetek ten wynosił 90%. W grupie dorosłych 77% wszystkich zapisów zostało zaklasyfikowanych w sposób identyczny pod kątem obecności odpowiedzi przez operatora i system automatyczny, a w grupie dzieci zgodność ta wyniosła 75%.

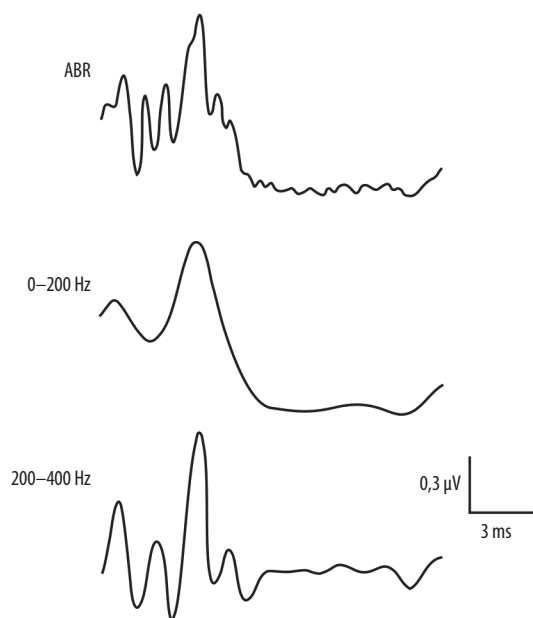


**Rycina 1.** Uśredniona odpowiedź ABR dla grupy 35 niemowląt oraz otrzymany przy jej wykorzystaniu wzorzec. Długość strzałek symbolizuje wagę dziewięciu punktów wzorca. Zgodnie z Killeny [15]

**Figure 1.** Grand average ABR for 35 infants and template derived from it. The lengths of arrows represent assigned weights. As in Killeny [15]

W 1987 r. Killeny [15] zaproponował system automatycznego wykrywania odpowiedzi oparty na rozpoznawaniu wzorców, wykorzystujący jako bodźce trzaski o natężeniu 35 dB nHL, filtrowane w taki sposób, aby ich widmo było płaskie w zakresie 750–5000 Hz. Wzorzec otrzymano, wykorzystując zapis poprawnej morfologii, powstały poprzez nałożenie na siebie 35 odpowiedzi dla trzasku 35 dB, zarejestrowanych u niemowląt ze słuchem normalnym. Wyznaczono dziewięć punktów w zapisie czasowym odpowiedzi, którym przypisano wagi w zależności od ich wkładu w detekcję odpowiedzi. Rycina 1 przedstawia schemat utworzenia wzorca. Fala V i następujące po niej obniżenie (SN10) uzyskały względnie wysokie wagi. Zastosowanie systemu dziewięciu ważonych punktów miało na celu zwiększenie efektywności algorytmu. Dopuszczalne, normalne przesunięcia latencji odpowiedzi zostały uwzględnione poprzez możliwość przesuwania bodźca w zakresie 3 ms w krokach 0,25 ms. Procedury przesuwania wzorca i zastosowania algorytmu były powtarzane co każde 500 pojedynczych rejestracji, aż do momentu uzyskania statystycznie istotnej odpowiedzi lub osiągnięcia maksymalnej dopuszczalnej liczby pojedynczych rejestracji, która została określona na 15 000. Co 500 rejestracji uśredniona odpowiedź była porównywana z wzorcem i wykonywano test statystyczny, czy przeprowadzono odpowiednią liczbę rejestracji do odróżnienia z 99,8-proc. przedziałem ufności sytuacji obecności odpowiedzi i szumu (wynik 'PASS') od sytuacji braku odpowiedzi lub obecności wyłącznie szumu. Przy osiągnięciu 15 000 powtórzeń i braku odpowiedzi system wyświetlał wynik 'REFER'.

Badania wstępne systemu przeprowadzono z udziałem 185 niemowląt (369 uszu). Algorytm osiągnął 94-proc. zgodność z osobami przeprowadzającymi badania. W pracy [16] system ten przetestowany został na większej grupie 231 niemowląt z wysokim ryzykiem występowania ubytku słuchu. W przypadku tych danych algorytm osiągnął czułość 89,2% i specyficzność 95,6% w porównaniu z konwencjonalnym badaniem ABR (Bio-logics LT). Zgodność wyniosła 95,3%. System ten został z założenia opracowany jako metoda wykrywania odpowiedzi ABR u niemowląt i zastosowany w urządzeniu ALGO-1.

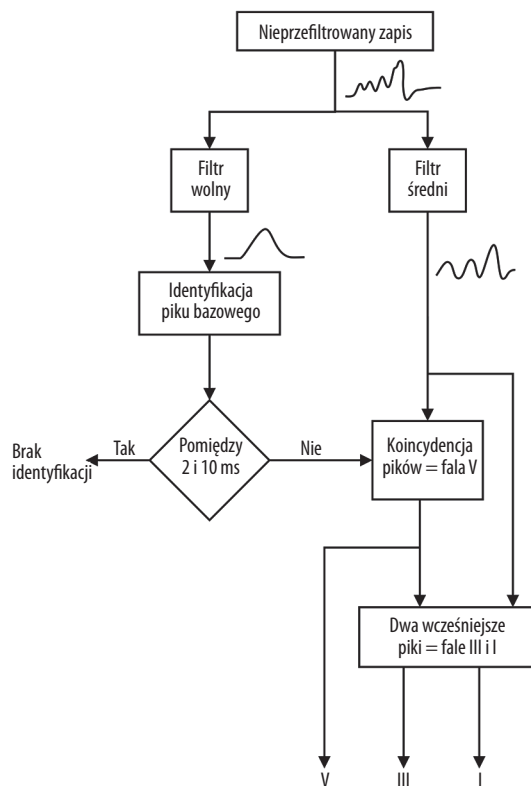


**Rycina 2.** Przykład zapisu ABR oraz wynik przefiltrowania go w dwóch pasmach do przebiegów czasowych, w których identyfikowane są pik bazowy oraz fale: I, III i V. Zgodnie z Pratt i wsp. [18]

**Figure 2.** An example of unfiltered ABR trace and waveforms obtained as results of filtering in two bands, where pedestal peak and waves: I, III and V can be identified. As in Pratt et al. [18]

W 1989 r. Pool i Finitzo zaproponowali algorytm detekcji fal I, III i V [17]. System ten oparty był na kryteriach dotyczących morfologii składowych odpowiedzi i stanowił próbę odwzorowania – za pomocą matematycznych operacji – zasad, jakimi kierują się doświadczeni klinicyści przy wzrokowej ocenie wyników badań ABR. Poszczególne bloki decyzyjne były warunkami obecności określonych struktur lub kryteriami dotyczącymi wartości parametrów takich jak amplituda czy latencja. W przypadku fal I i III algorytm był taki sam jak dla fali V, tylko ze zmienionymi wartościami parametrów. W celu oszacowania wydajności algorytmu wyniki jego działania zestawiono z wynikami ocen czterech techników doświadczonych w badaniach ABR. Grupę testową stanowiły osoby ze słuchem normalnym i z ubytkami słuchu, w wieku od 6 miesięcy do 75 lat. W sumie przeanalizowano 248 zapisów ABR zarejestrowanych przy minimalnej liczbie 2000 uśrednień dla natężeń bodźca w zakresie 20–103 dB HL. Rezultaty ocen zarówno czterech techników, jak i systemu zostały ocenione przez dwóch ekspertów. System automatyczny nie odstawał w wynikach ocen od techników. Na 248 zapisów algorytm miał 2 błędy, ale trzech z czterech techników również miało przynajmniej dwa błędy. Technicy mieli wyłącznie błędy fałszywie dodatnie, przejawiając tendencję do doszukiwania się odpowiedzi tam, gdzie według ekspertów jej nie było. System natomiast miał jeden błąd fałszywie dodatni i jeden fałszywie ujemny.

W innej pracy z 1989 r. Pratt z zespołem opisali przykład zastosowania częstotliwościowych metod analizy sygnałów



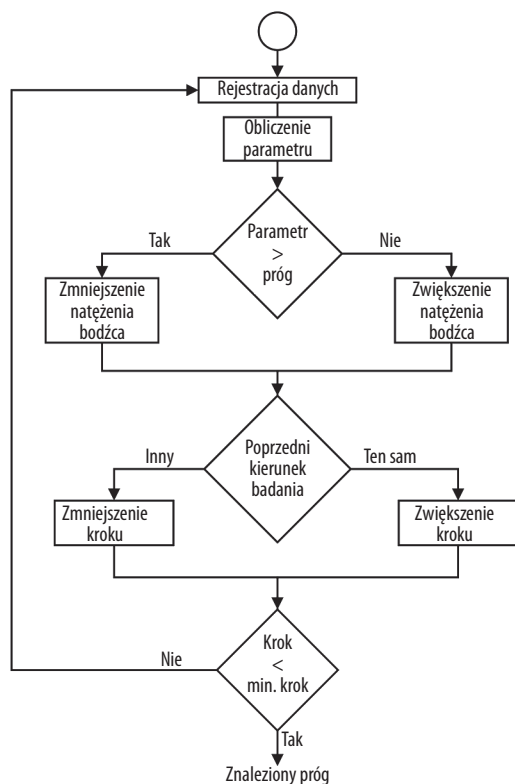
**Rycina 3.** Schematyczne przedstawienie procedury rozpoznania składowych fal V, III i I. Zgodnie z Pratt i wsp. [18]

**Figure 3.** Schematic representation of identification of wave V, III and I. As in Pratt et al. [18]

do automatycznej detekcji fali V [18]. Badając widmo ABR, autorzy wyodrębnili trzy pasma częstotliwości: 0–240 Hz, 240–483 Hz i powyżej 500 Hz. W pierwszym paśmie (0–240 Hz) daje się wyróżnić tzw. pik bazowy, a fale: I, III i V są widoczne w paśmie drugim (240–483 Hz). Na podstawie tych obserwacji działanie systemu wyglądało następująco. Po przefiltrowaniu sygnału do określonych pasm identyfikowano pik bazowy z pierwszego pasma i wszystkie piki z pasma drugiego. Po sprawdzeniu warunku, czy latencja pików bazowych mieści się między 2 i 10 ms zapisu, fala V była oznaczana jako pik z drugiego pasma koincydujący w latencji z pikiem bazowym. Fale I i III identyfikowane były jako dwa wcześniejsze piki w drugim paśmie. Na rycinie 2 przedstawiono przykładowy zapis ABR oraz wynik jego przefiltrowania i rozpoznania fal. Rycina 3 przedstawia schemat procedury identyfikacji składowych. System wykorzystano do oceny badań ABR w grupie ochotników. Przebadano 23 uszu u 14 osób ze słuchem w normie. Odpowiedzi były rejestrowane dla trzasku, 75 dB nHL, z częstotnością podawania bodźca 10, 55 i 80 Hz oraz dla trzasku 45 i 15 dB nHL, z częstotnością 10 Hz, z polaryzacją naprzemienną w każdym przypadku. Autorzy ocenili algorytm jako wiarygodny we wszystkich próbach i jako posiadający bardzo dobrą zgodność z oceną wzrokową.

Praca [19] z 1994 roku jest przykładem zastosowania adapttywnej procedury szukania progu (ang. *Parameter Estimation by Sequential Testing*, PEST) realizowanej w trakcie badania i sterowanej modulem rozpoznającym odpowiedź



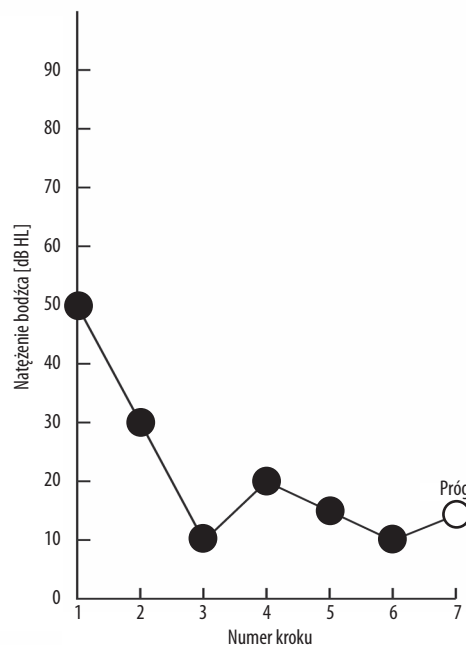


**Rycina 4.** Schemat algorytmu PEST określenia progu odpowiedzi. Zgodnie z Ozdamar i wsp. [19]

**Figure 4.** PEST algorithm for response threshold identification. As in Ozdamar et al. [19]

opartym na korelacji. Rejestracja pojedynczych odpowiedzi odbywała się w dwóch buforach. Część sygnału, w której należy oczekiwać najsilniejszej odpowiedzi, dzielona była na 7 krótkich okien po 2 ms, w których obliczane są korelacje pomiędzy zapisami w obu buforach. Graniczna wartość współczynnika korelacji została zoptymalizowana i wynosiła 0,7. Początkowa wielkość zmiany natężenia bodźca to 20 dB. W kolejnych krokach algorytmu – w przypadku zmiany obecności odpowiedzi w zapisie – wartość zmiany natężenia była zmniejszana o połowę, a przyrost był zmieniany na przeciwny. Procedurę powtarzano aż do momentu ustalenia progu ABR. Schemat algorytmu PEST jest przedstawiony na rycinie 4, a przykładowy przebieg badania z użyciem tego systemu na rycinie 5.

Dane, na których testowany był system, pochodziły od osób z ubytkami słuchu i osób ze słuchem w normie, zarówno dorosłych, jak i dzieci. Grupa dorosłych liczyła 66 osób (28 mężczyzn, 38 kobiet) w wieku 13–85 lat, w sumie 131 zapisów. Grupę niemowląt i małych dzieci stanowiło 39 osób (26 chłopców, 13 dziewczynek) w wieku od 11 tygodni do 3,6 roku, w sumie 72 zapisy. Czas trwania większości badań progowych nie przekraczał 15 min. Wyniki działania systemu porównano z oznaczeniami ekspertów. Algorytm osiągnął dużą zgodność w oznaczeniu progu ABR z ekspertami, zarówno w grupie dorosłych jak i dzieci. Jednak porównanie otrzymanych automatycznie progów ABR z wynikami audiometrii tonalnej dało w grupie dorosłych duży rozrzut zarówno dla niskich, jak



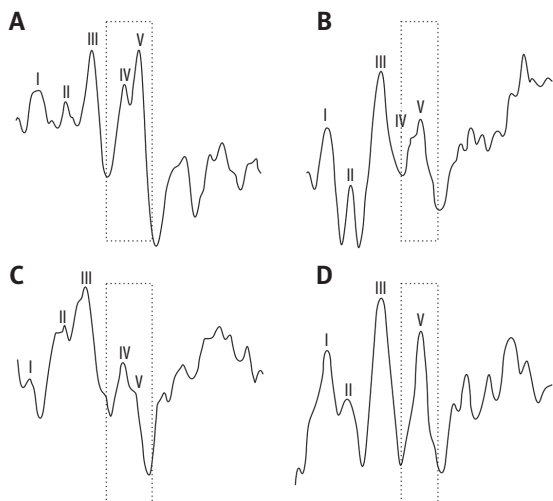
**Rycina 5.** Przykładowy przebieg badania ABR przy zastosowaniu procedury PEST. Zgodnie z Ozdamar i wsp. [19]

**Figure 5.** An example of the PEST procedure in ABR threshold testing. As in Ozdamar et al. [19]

i wysokich częstotliwości audiometrycznych. Pojedyncze punkty odstawały od linii trendu, jednak zachowana była ogólna zgodność system–ekspert. Wyniki określenia progu przez system automatyczny były bardziej zbliżone do oznaczeń ekspertów w grupie dzieci niż w grupie dorosłych w przypadku niskich wartości progu ABR. W przypadku wyższych wartości progu rozrzut w oznaczeniach progu system–ekspert był większy w grupie dzieci niż dorosłych.

W 1999 r. Zajac wraz z zespołem [20] zaproponowali system detekcji fali V bazujący na funkcji korelacji zapisów z wzorcami odpowiedzi. Algorytm działał dla odpowiedzi na bodźce typu trzask o natężeniu 10–100 dB nHL, rejestrowanych metodą szeregu natężeniowego. Początkowo testowane wzorce, utworzone na podstawie uśrednionych zapisów 50 osób o słuchu normalnym, nie dawały satysfakcjonujących efektów ze względu na różnorodność morfologii odpowiedzi, w szczególności dla natężeń okołoprogowych. Dopiero wzorce wytworzone na podstawie wzorca fali V dla bodźca o natężeniu 90 dB nHL dawały satysfakcjonujące rezultaty. Latencja fali V określana była przez system zgodnie z lokalizacją czasową maksimum funkcji korelacji wzorca z zapisem, przy uwzględnieniu dwóch warunków: na minimalną wartość maksimum funkcji korelacji ( $650 \text{ nV}^2$ ) i jej gradient (osiągnięcie 75% maksymalnej wartości w przeciągu 1 ms). System przetestowano na zapisach zarejestrowanych w grupie 30 osób o słuchu w normie i 70 osób z ubytkami słuchu typu ślimakowego (w sumie 304 odpowiedzi). Porównanie wyników z oznaczeniami zespołu ekspertów pokazało, że algorytm miał 93-proc. zgodność z oznaczeniami ekspertów.

Popescu i wsp. w 1999 r. zaproponowali algorytm detekcji fali V [21]. Metoda ta złożona była z trzech modułów.



**Rycina 6.** Cztery możliwe klasyfikacje odpowiedzi (A, B, C i D) ze względu na morfologię kompleksu fal IV i V. Zgodnie z Popescu i wsp. [21]

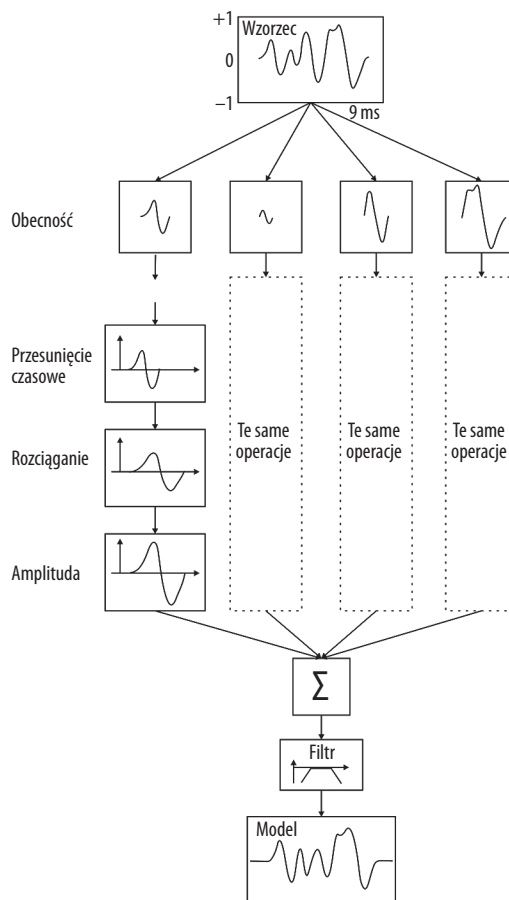
**Figure 6.** Four different classes (A, B, C, and D) based on different morphologies of the IV–V complexes. As in Popescu et al. [21]

W pierwszym module zarejestrowany sygnał był poddawany dekompozycji za pomocą transformacji falkowej. W drugim kroku z zapisu usuwano szum przy wykorzystaniu trójwarstwowej sieci neuronowej. Filtrowanie to oparte było na fakcie, że składowe odpowiedzi są reprezentowane przez falki o innych właściwościach niż składowe szumu. W ostatnim kroku działania systemu wykrywane były składowe fale. Ta część opierała się na kryteriach. Fale były identyfikowane jako jedna z czterech możliwych klas odpowiedzi A, B, C lub D, przedstawionych na rycinie 6, na podstawie morfologii kompleksu fal IV i V. Klasa A – dotyczyła sytuacji, kiedy w zapisie dają się wyróżnić fala IV i fala V jako dwa oddzielne szczyty, klasa B – kiedy fala IV jest punktem przegięcia na lewym zboczu wstępującym szczytu fali V, klasa C – kiedy fala V jest punktem przegięcia na prawym zboczu zstępującym szczytu fali IV, i klasa D – gdy w zapisie widoczna jest wyłącznie fala V.

Dzięki temu, że analiza ABR przebiega w przypadku tego systemu w wielu skalach jednocześnie, gdyż różne częstotliwości są osiągane poprzez rozciąganie falki bazowej, metoda ta nie jest obciążona stratą informacji związaną z filtrowaniem sygnału w określonych pasmach częstotliwości. Większość składowych szumu jest usuwana przez sieć neuronową w fazie drugiej, więc nie wywierają one wpływu na jakość modułu rozpoznawania struktur, dzięki czemu metoda jest odporna na zakłócenia szumowe.

W celu oszacowania wydajności algorytmu przebadano grupę 52 osób z prawidłowymi (38 osób) i nieprawidłowymi (2 osoby z nowotworem pnia mózgu, 8 osób z przewodzeniowymi ubytkami słuchu) zapisami ABR. W grupie było czworo dzieci w wieku 2–4 lat. System dobrze rozpoznawał odpowiedź w badanych grupach praktycznie dla każdej z czterech możliwych postaci odpowiedzi (A, B, C i D).

W pracy [22] zaproponowano system identyfikacji składowych odpowiedzi na podstawie dekompozycji modelu

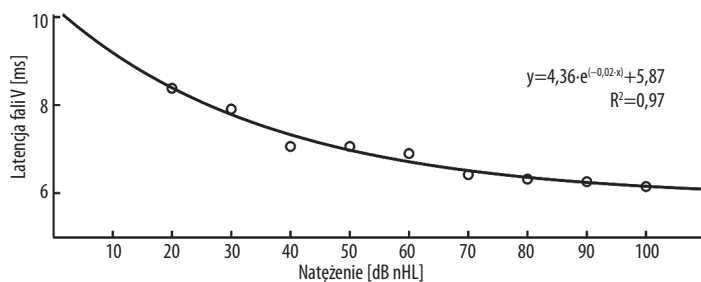


**Rycina 7.** Schemat tworzenia matematycznego modelu odpowiedzi z dopasowanych, przesuniętych w czasie i przeskalowanych w amplitudzie fragmentów wzorców składowych. Zgodnie z Vannier i wsp. [24]

**Figure 7.** Mathematical model of ABR is created from time shifted, stretched and amplitude adjusted elements of template. As in Vannier et al. [24]

zapisu stworzonego z zastosowaniem wzorców. Trzeci współautor publikacji – Motsch w pracy z 1987 r. [23] przedstawił cztery wzorce odpowiedzi wystarczające do poprawnego modelowania zapisu ABR. W algorytmie z 2000 r. wzorce Motscha dzielone były na cztery fragmenty, a spośród nich wybierano najlepiej pasujące do zapisu. Po dopasowaniu amplitudy i przesunięcia fragmentów tworzony był przy ich użyciu optymalny model odpowiedzi. Po znalezieniu najlepszego modelu dalsza analiza wykonywana była nie na oryginalnym sygnale, ale na jego modelu. Latencje składowych sygnału identyfikowano z zastosowaniem transformacji falkowej. Po przetestowaniu na zapisach z grupy pacjentów ze słuchem w normie (40% osób) i z ubytkami słuchu (54% – pacjenci z ubytkami ślimakowymi i 6% z ubytkami pozaślimakowymi) autorzy podali, że system osiągnął 83-proc. poprawność detekcji (w przypadku falki Symlet).

W 2001 r. Vannier i inni przedstawili system automatycznej detekcji fali V oraz fal I i III [24,25]. Do rozpoznawania składowych wykorzystano system wzorców Motscha



**Rycina 8.** Przykładowa krzywa latencja-natężenie otrzymana dla osoby ze słuchem w normie oraz równanie będące wynikiem dopasowania do niej funkcji wykładniczej o podstawie e (podstawie logarytmu naturalnego)

**Figure 8.** An example of latency-intensity curve obtained for a person with normal hearing and the equation of exponential function fitted to data

– wówczas system działał w trybie automatycznej detekcji nienadzorowanej. Matematyczny model ABR otrzymywany był na podstawie dopasowanych i przeskalowanych wzorców składowych fal, tak jak pokazano na rycinie 7.

Nową rzeczą w pracy było uwzględnienie w procesie rozpoznawania składowych zależności latencji fali V od natężenia bodźca – automatyczna detekcja nadzorowana. Zależność ta ma postać charakterystycznej krzywej (ang. *Latency-Intensity curve*, LI). Na rycinie 8 przedstawiona jest przykładowa krzywa latencja-natężenie oraz wynik dopasowania do niej funkcji eksponencjalnej. Autorzy przebadali kilka klas krzywych, wyznaczając równanie najlepiej pasującej funkcji. Matematycznym modelem najlepiej pasującym do danych była funkcja odwrotności trójmianu kwadratowego. Ostatecznie jednak autorzy rozbili zależność na dwie części: dla wysokich (H) i niskich (L) natężeń bodźca, przyjmując jako punkt graniczny wartość średniego progu audiometrycznego w przedziale częstotliwości 2–4 kHz powiększoną o 40 dB. Informacja kontekstowa, wynikająca z uwzględnienia tej zależności, była wykorzystywana do nadzorowania procesu rozpoznawania wzorców. Fale I i III wykrywane były w sposób analogiczny.

Dokładność automatycznego tworzenia krzywej latencja-natężenie przetestowano na 190 badaniach przeprowadzonych metodą szeregu natężeniowego (190 krzywych LI). Wyniki porównano z oceną ekspercką. Dla fali V system był zgodny w ocenie w 83% zapisów, niezgodny w 3%. W 9% na płaszczyźnie LI występowała zbyt mała liczba punktów, w 1% musiała zostać wykorzystana metoda ręczna, a w 4% interpretacja zapisu ABR była niemożliwa. Dla fal III i I system poprawnie wykrył odpowiednio: 90% i 76,5% przypadków. Wartościami tolerancji przy porównaniu działania systemu z ekspertem była wartość 0,2 ms w latencji fali i 10 dB w przypadku określenia progu odpowiedzi ABR. Średnia różnica pomiędzy systemem a ekspertem w określeniu latencji fali V wynosiła 0,05 ms, a w wartości progu ABR 5 dB.

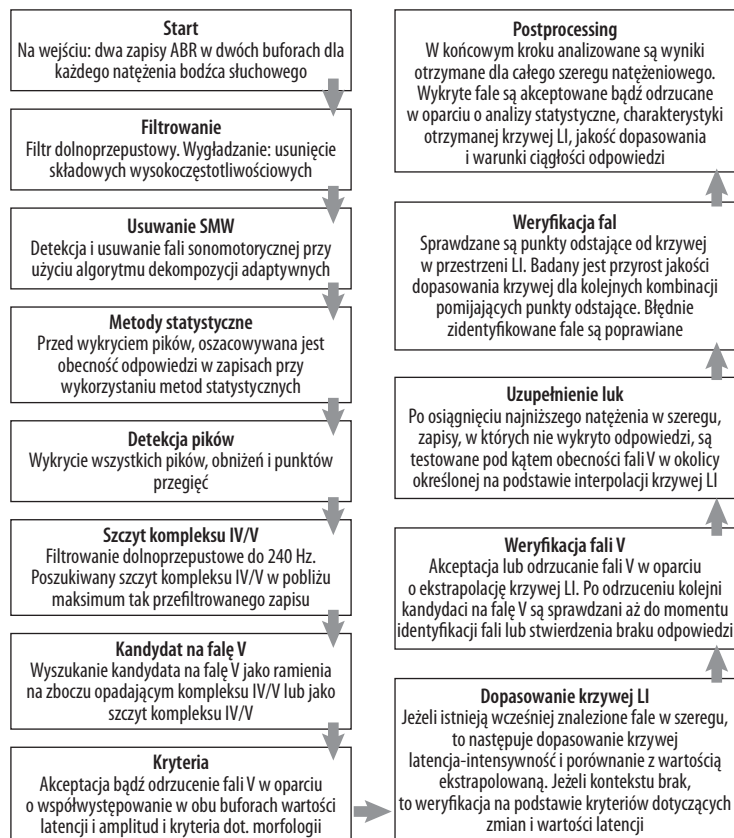
System przetestowano na 283 zapisach ABR dla 20 losowo wybranych pacjentów. Dla fali V system nadzorowany (uwzględniający informację z krzywej LI) osiągnął czułość 91% i specyficzność 92%. Algorytm przebadano też w pracy klinicznej na 30 uszach (213 zapisów ABR). Uzyskano dla fali V czułość 89% i specyficzność 92%, a po odrzuceniu zapisów problematycznych w interpretacji: czułość 99% i specyficzność 92%.

Praca z 2005 r. to przykład metody opartej w zdecydowanej mierze na kryteriach i warunkach [26]. System wykrywał

fale I–VII w normalnych zapisach ABR dla bodźców ponadprogowych. Algorytm działał w trzech lub opcjonalnie w czterech etapach. W pierwszym kroku, przy wykryciu miejsc zerowych pierwszej pochodnej sygnału, wykrywano wszystkie lokalne maksima i minima. W drugim etapie fale I–VII rozpoznawano na podstawie kryteriów na latencję i morfologię składowych. Jeżeli zachodziła taka potrzeba, w trzecim kroku fale II i IV wykrywano przy użyciu miejsc zerowych drugiej pochodnej zapisu. Na ostatnim etapie, w celu dokładnego określenia latencji i amplitud fal, identyfikowano obniżenia pomiędzy składowymi falami.

W celu oszacowania wydajności algorytmu wyniki automatyczne zostały porównane z oznaczeniami ekspertów dla 240 zapisów ABR zarejestrowanych w grupie 60 osób ze słuchem normalnym (22 mężczyzn i 38 kobiet), w wieku od 18 do 55 lat. Wszystkie rejestracje wykonane były dla bodźca typu trzask o natężeniu 90 dB nHL. System poprawnie zidentyfikował trzy główne fale (I, III i V) w 96–98% przypadków. Dla fali V system osiągnął maksymalną wartość czułości 98%. Wydajność systemu w przypadku pozostałych fal (II, IV, VI i VII) wahała się w przedziale 45–83% poprawnych detekcji. W przypadku fal II i IV czułość metody wzrosła odpowiednio do: 83% i 77%, przy dopuszczeniu możliwości oznaczenia fal jako punktów przegięcia w zapisie ABR. Autorzy zasugerowali, że podobna klasyfikacja musi być podświadomie przeprowadzana przez eksperta oceniającego badania w przypadku braku tych fal.

W 2007 r. opublikowana została statystyczna metoda detekcji ABR wykorzystująca stale rosnącą moc obliczeniową komputerów [27]. Metoda ta oparta była na technice bootstrapu polegającej na wielokrotnym losowym repróbowaniu ze zwracaniem danych oryginalnych, co dawało w wyniku estymatę prawdopodobieństwa, że zmierzone dane są wynikiem statystycznej fluktuacji, a nie zarejestrowania rzeczywistej odpowiedzi elektrofizjologicznej. Metoda ta pozwalała oszacować istotność statystyczną obliczonego na sygnale parametru. Autorzy przebadali skuteczność metody dla czterech parametrów: różnicy pomiędzy maksymalną a minimalną wartością zapisu ABR, energii sygnału w przedziale czasowym 5–15 ms od momentu podania bodźca,  $F_{sp}$  i alternatywnej estymaty stosunku sygnału do szumu. Metoda oparta na technice bootstrapu jest w rzeczywistości całą klasą metod o wydajności zależnej od parametru, którego istotność szacuje. Ponieważ w wyniku otrzymywane jest prawdopodobieństwo, że obliczona wartość parametru jest wynikiem obecności odpowiedzi w zapisie, dobrze nadaje się do zastosowań takich jak np. przesiewowe badania ABR.



**Rycina 9.** Schemat blokowy działania systemu detekcji fali V. Zgodnie z Trzaskowski i wsp. [28]

**Figure 9.** Block diagram of wave V identification system. As in Trzaskowski et al. [28]

Algorytm przedstawiony w [28,29] działał dla standardowych bodźców typu trzask 100  $\mu$ s, wykrywając próg ABR i latencje fal V. System do pracy potrzebował zapisów z dwóch buforów lub dwóch uśrednionych sygnałów. Na wstępie działania systemu zapisy były filtrowane i wygładzane. Jeżeli zawierały falę somomotoryczną, to była ona automatycznie rozpoznawana i usuwana. Zanim algorytm przystąpił do detekcji fali V, najpierw za pomocą metod statystycznych określał, czy w sygnale występuje odpowiedź. Zaimplementowane metody obejmowały między innymi metody korelacyjne czy bootstrap z różnymi parametrami. W kolejnym kroku wykrywano wszystkie piki i za pomocą mocnego filtrowania analogicznego do metody piku bazowego [18] identyfikowano był kompleks fal IV/V. W kolejnych krokach – przy wykorzystaniu kryteriów dotyczących latencji, amplitudy i morfologii oraz przede wszystkim dopasowania krzywej latencja-intensywność – znajdowana była fala V. Procedurę powtarzano do momentu znalezienia progu odpowiedzi. Na sam koniec, na podstawie całościowej informacji kontekstowej z krzywej latencja-intensywność, uzupełniano niewykryte fale V i poprawiano fale wykryte błędnie. Zasadę działania algorytmu schematycznie przedstawiono na rycinie 9.

Algorytm został przetestowany na danych klinicznych w grupie 21 osób ze słuchem normalnym i 14 pacjentów z ubytkami słuchu, w wieku 19–68 lat, a wyniki jego działania zostały porównane z ocenami ekspertów. Na zbiorze 35 badań progowych system osiągnął 94-proc. zgodność określenia progu z oznaczeniami ekspertów w zakresie 10 dB.

Określenia latencji fali V wykonane przez system automatyczny dobrze korelowały z oznaczeniami ekspertów.

Algorytm został także przetestowany pod kątem efektywności w oznaczaniu progu odpowiedzi na dużej liczbie danych klinicznych. Przeanalizowano 123 szeregi natężeniowe (740 zapisów) w grupie osób z prawidłowym słuchem i 246 szeregow natężeniowych (1292 zapisy) w grupie osób z ubytkami słuchu typu ślimakowego. Na tym zbiorze danych system osiągnął 98-proc. zgodność z ekspertami w oznaczeniu progu w przedziale do 20 dB w grupie osób ze słuchem normalnym i 91-proc. zgodność w grupie pacjentów z ubytkami słuchu. Dla obu grup różnice do 30 dB uzyskano w 99% przypadków.

## Podsumowanie

Słuchowe potencjały wywołane pnia mózgu są jedną z najczęściej stosowanych i zarazem najlepiej poznanych metod badania słuchu. Zaliczane są do grupy badań obiektywnych, ponieważ nie wymagają współpracy ze strony osoby badanej, jednak w przypadku części testów, jak np. badanie progowe lub diagnostyka różnicowa ABR, wynik wciąż oceniany jest wzrokowo przez osobę wykonującą badanie.

Od momentu odkrycia tych potencjałów w 1970 r. trwają prace nad zwiększeniem obiektywizacji i standaryzacji oceny wyników badań. W pracy podjęto próbę scharakteryzowania różnych podejść do problemu automatycznej detekcji ABR. Przedstawiono wybrane systemy bazujące na różnych technikach oraz wyniki ich zastosowania w pracy klinicznej.



Na obecnym etapie obiecujące wydają się techniki wykorzystujące dużą moc obliczeniową współczesnych komputerów, jak np. przedstawiony w pracy [27] system wykorzystujący technikę bootstrapu. Ta technika ma tę zaletę, że pozwala oszacować istotność odpowiedzi ABR na bodziec słuchowy względem tego samego zapisu, ale jakby rejestrowanego bez bodźca. Jest to możliwe dzięki temu, że repróbkiwanie danych pozwala zlikwidować zależność czasową pojedynczych rejestracji od pobudzenia. Porównanie wartości parametru dla odpowiedzi na bodziec z wartościami parametru dla danych reprobkiwanych ze skasowaną zależnością czasową pozwala oszacować istotność odpowiedzi ABR względem zapisu bez bodźca dokładnie w tym samym stanie.

Jednocześnie sama metoda reprobkiwania stanowi jedynie mechanizm umożliwiający oszacowanie istotności statystycznej wykrytej odpowiedzi. Wydajność całego systemu zależy od zastosowanego parametru bootstrapu. Oznacza to, że teoretycznie możliwe jest stworzenie systemu o dowolnej efektywności dzięki wykorzystaniu odpowiedniego parametru. To powoduje, że metody tego typu mają duży potencjał rozwoju.

Obecnie ograniczeniem może być to, że systemy oparte na reprobkiwaniu wymagają znacznej mocy obliczeniowej. Koszt obliczeniowy rośnie liniowo wraz z liczbą powtórzeń bootstrapu. Dodatkowo złożoność przekształceń zależy od komplikacji obliczeniowej parametru bootstrapu i dla skomplikowanego numerycznie parametru koszt obliczeniowy metody może być bardzo wysoki. Jednakże stale rosnąca wydajność komputerów oraz możliwość wykorzystania w obliczeniach w coraz większym stopniu procesorów graficznych pozwalają przypuszczać, że w przyszłości będą mogły być stosowane metody coraz bardziej kosztowne obliczeniowo.

Pomimo istnienia wielu opisanych systemów wciąż brakuje jednego powszechnie akceptowanego systemu automatycznej detekcji słuchowych potencjałów wywołanych pnia mózgu. W przypadku większości badań ABR człowiek nadal odgrywa decydującą rolę w procesie oceny wyniku badania.

*Publikacja powstała w związku z realizacją projektu pn. „Zintegrowany system narzędzi do diagnostyki i telerehabilitacji schorzeń narządów zmysłów (słuchu, wzroku, mowy, równowagi, smaku, powonienia)” INNOSENSE, współfinansowanego przez Narodowe Centrum Badań i Rozwoju w ramach Programu STRATEGMED.*

## Piśmiennictwo:

- Jewett DL, Romano MN, Williston JS. Human auditory evoked potentials: possible brain stem components detected on the scalp. *Science*, 1970; 13: 1517–8.
- Jewett D, Williston J. Auditory-evoked far fields averaged from the scalp of humans. *Brain*, 1971; 94: 681–96.
- Katz J, Burkard R, Medwetsky L. *Handbook of Clinical Audiology*. Fifth Edition. Lippincott, Williams, and Wilkins; 2002.
- Trzaskowski B, Jedrzejczak WW, Pilka E, Kochanek K, Skarżyński H. Automatic removal of somatomotor waves from auditory brainstem responses. *Comput Biol Med*, 2013; 43(5): 524–32.
- Vidler M, Parkert D. Auditory brainstem response threshold estimation: subjective threshold estimation by experienced clinicians in a computer simulation of the clinical test. *Int J Audiol*, 2004; 43: 417–29.
- Elberling C. Auditory electrophysiology. The use of templates and cross correlation functions in the analysis of brain stem potentials. *Scand Audiol*, 1979; 8(3): 187–90.
- Weber BA, Fletcher GL. A computerized scoring procedure for auditory brainstem response audiometry. *Ear Hear*, 1980; 1(5): 233–6.
- Elberling C, Don M. Quality estimation of averaged auditory brainstem responses. *Scand Audiol*, 1984; 13(3): 187–97.
- Don M, Elberling C, Waring M. Objective detection of averaged auditory brainstem responses. *Scand Audiol*, 1984; 13(4): 219–28.
- Schimmel H. The (+/-) reference: accuracy of estimated mean components in average response studies. *Science*, 1967; 157(3784): 92–4.
- Wong PK, Bickford RG. Brain stem auditory evoked potentials: the use of noise estimate. *Electroencephalogr Clin Neurophysiol*, 1980; 50(1–2): 25–34.
- Mason SM. On-line computer scoring of the auditory brainstem response for estimation of hearing threshold. *Audiology*, 1984; 23(3): 277–96.
- Mason SM. Objective waveform detection in electric response audiometry. *Electroencephalogr Clin Neurophysiol*, 1985; 61(3): 242.
- Mason SM. Automated system for screening hearing using the auditory brainstem response. *Br J Audiol*, 1988; 22(3): 211–3.
- Kileny PR. Algo-1 automated infant hearing screener: preliminary results. *Semin Hear*, 1987; 8: 125–31.
- Jacobson JT, Jacobson CA, Spahr RC. Automated and conventional ABR screening techniques in high-risk infants. *J Am Acad Audiol*, 1990; 1(4): 187–95.
- Pool KD, Finitzo T. Evaluation of a computer-automated program for clinical assessment of the auditory brain stem response. *Ear Hear*, 1989; 10(5): 304–10.
- Pratt H, Urbach D, Bleich N. Auditory brainstem evoked potentials peak identification by finite impulse response digital filters. *Audiology*, 1989; 28(5): 272–83.
- Ozdamar O, Delgado RE, Eilers RE, Urbano RC. Automated electrophysiologic hearing testing using a threshold-seeking algorithm. *J Am Acad Audiol*, 1994; 5(2): 77–88.
- Zajac J, Kochanek K, Pietraszek S, Piłka A, Skarżyński H. Automatyczne oznaczanie szczytu fali V słuchowych potencjałów wywołanych pnia mózgu za pomocą funkcji korelacji – wyniki wstępne. *Audiofonologia*, 1999; XV: 221–30.
- Popescu M, Papadimitriou S, Karamitsos D, Bezerianos A. Adaptive denoising and multiscale detection of the V wave in brainstem auditory evoked potentials. *Audiol Neurootol*, 1999; 4(1): 38–50.
- Naït-Ali A, Adam O, Motsch JF. Modelling and recognition of brainstem auditory evoked potentials using Symlet wavelet. *ITBM-RBM*, 2000; 21(3): 150–7.
- Motsch JF. La dynamique du tronc cerebral [these]. Université Paris, 1987.

24. Vannier E, Adam O, Karasinski P, Ohresser M, Motsch JF. Computer-assisted ABR interpretation using the automatic construction of the latency-intensity curve. *Audiology*, 2001; 40(4): 191–201.
25. Vannier E, Adam O, Motsch JF. Objective detection of brainstem auditory evoked potentials with a priori information from higher presentation levels. *Artif Intell Med*, 2002; 25(3): 283–301.
26. Bradley AP, Wilson WJ. Automated analysis of the auditory brainstem response using derivative estimation wavelets. *Audiol Neurootol*, 2005; 10(1): 6–21.
27. Lv J, Simpson DM, Bell SL. Objective detection of evoked potentials using a bootstrap technique. *Med Eng Phys*, 2007; 29(2): 191–8.
28. Trzaskowski B, Jędrzejczak W, Piłka E, Kochanek K, Skarżyński H. System automatycznej detekcji słuchowych potencjałów wywołanych pnia mózgu. I. Opis i testowanie systemu. *Otorynolaryngologia*, 2013; 12(3): 137–47.
29. Trzaskowski B, Kochanek K, Jędrzejczak W, Piłka A, Skarżyński H. System automatycznej detekcji słuchowych potencjałów wywołanych pnia mózgu. II. Ocena działania systemu dla danych klinicznych. *Otorynolaryngologia*, 2013; 12(4): 183–9.