Statistics in medical research

Wojciech Jasek, Włodzimierz Stec

Clinical Department of Reconstructive Plastic Surgery and Burn Treatment, Military Medical Institute, Warsaw

Corresponding author: Dr. Wojciech Jasek, E-mail: wojciechjas@o2.pl

ABSTRACT

Every doctor who in his professional career wants to share his experiences in the pages of specialized medical journals, faces the problem of analysing and evaluating his material in a statistical manner. Given the specificity of our profession, it is not easy. In order to meet this problem, the paper set the goal to briefly systematize and explain the basic statistical methods needed in publishing scientific papers by doctors. The method of selecting the research sample is discussed and the necessary concepts necessary for its implementation are explained. It was divided into a random and non-random sample. The theory and method of application of measurement scales were briefly discussed and the practical division of these scales was made. The concise content and form allows for independent use of statistical methods in the basic scope needed in publishing medical works.

Key words: Sample selection; Research sample; Sample representativeness; Measurement scales

How to cite this article: Jasek W, Stec W. Statistics in medical research. Our Dermatol Online. 2018;9(e):e6.1-e6.5.

Submission: 15.11.2018; **Acceptance:** 02.12.2018

DOI:10.7241/ourd.2018e.6

Statystyka w badaniach medycznych

Wojciech Jasek, Włodzimierz Stec

Clinical Department of Reconstructive Plastic Surgery and Burn Treatment, Military Medical Institute, Warsaw

Corresponding author: Dr. Wojciech Jasek, E-mail: wojciechjas@o2.pl

STRESZCZENIE

Każdy lekarz, który w swojej pracy zawodowej chce podzielić się swoimi doświadczeniami na łamach specjalistycznych czasopism medycznych, staje przed problemem analizy i oceny swojego materiału w sposób statystyczny. Biorąc pod uwagę specyfikę naszego zawodu nie jest to łatwe. Wychodząc naprzeciw temu problemowi, w pracy postawiono za cel krótkie usystematyzowanie i wyjaśnienie podstawowych metod statystycznych, potrzebnych w publikowaniu prac naukowych przez lekarzy. Omówiono sposób doboru próby badawczej i wyjaśniono potrzebne terminy, które są niezbędne do jej wykonania. Podzielono ją na próbę losową i nielosową. Krótko omówiono teorię i sposób zastosowania skal pomiarowych oraz dokonano praktycznego podziału tych skal. Zwięzła treść i forma pozwala na samodzielne korzystanie w podstawowym zakresie z metod statystycznych, potrzebnych w publikowaniu prac medycznych.

Słowa Klucze: Dobór próby; Próba badawcza; Reprezentatywność próby; Skale pomiarowe

WSTĘP

Każde badanie o charakterze naukowym powinno spełniać określone wymogi związane z ich przeprowadzaniem, dokumentowaniem oraz publikowaniem wyników. Realizując takie badania należy dołożyć wszelkich starań, aby były one rzetelne i zgodne z przyjętymi zasadami. Warto więc zadbać o prawidłowe ich przygotowanie, bo to pomaga uniknąć błędów, które mogą zniweczyć efekt podjętej pracy.

Metody reprezentacyjne na dobre upowszechniły się już w naukach społecznych. Badając grupę osób o relatywnie małej liczebności uzyskuje się wyniki, które są słuszne dla całej populacji. Najbardziej znanym wykorzystaniem metod probabilistycznych (losowych) jest badanie opinii publicznej w zakresie preferencji wyborczych oraz w celach marketingowych. W badaniach medycznych coraz częściej również sięga się po metody statystyczne, które ułatwiają wyciąganie wniosków z uzyskanych pomiarów i obserwacji. Możliwość wykorzystania metod statystycznych bardzo często związana jest z koniecznością spełnienia określonych wymogów [1-3]. Do najważniejszych

zaliczyć należy:

- wybór metody statystycznej możliwej do zastosowania ze względu na dane którymi się dysponuje,
- wybór skali pomiarowej, na której dokonuje się pomiarów,
- dobór wielkości próby badawczej (w statystyce przez próbę badawczą rozumie się grupę osób poddanych obserwacji (badaniu)),
- zapewnienie jej reprezentatywności.

Dobór Próby Badawczej

Wiele z badań naukowych polega na poddawaniu obserwacji grupy pacjentów (próby badawczej) o określonej liczebności, a uzyskane wyniki uogólniane są następnie na całą populację. Aby takie uogólnienie było słuszne, grupa (próba badawcza) poddana badaniu powinna być odpowiednio liczna, a ponadto musi odzwierciedlać cechy całej populacji dla której chcemy formułować wnioski, czyli być wobec niej reprezentatywną [1,2,4].

Istnieją dwie podstawowe metody doboru próby do badań:

- próba probabilistyczna (losowa),
- próba nieprobabilistyczna.

How to cite this article: Jasek W, Stec W. Statystyka w badaniach medycznych. Our Dermatol Online. 2018;9(e):e6.1-e6.5.

Submission: 15.11.2018; **Acceptance:** 02.12.2018

DOI:10.7241/ourd.2018e.6

Próba losowa dobierana jest w taki sposób, że dla każdej jednostki należącej do populacji:

- znane jest prawdopodobieństwo jej wylosowania do próby,
- prawdopodobieństwo to jest większe od zera.

Obecnie probabilistyczny (losowy) dobór grupy badanej jest podstawową metodą dobierania reprezentatywnych prób na potrzeby badań społecznych. Taki przypadek występuje również, kiedy badania mają na celu weryfikację stanu zdrowotności społeczeństwa lub skuteczności leku. W odniesieniu do tego typu badań występuje oczekiwanie, że wynik uzyskany dla próby badawczej będzie również prawdziwy w całej populacji, której to badanie dotyczy. Istnieje wiele metod doboru próby, które zapewniają losowy jej charakter [4,5].

W badaniach medycznych niejednokrotnie celem prowadzonych analiz jest stwierdzenie skuteczności lub przydatności wykonywanego zabiegu lub procedury medycznej. Przy tego typu badaniach nie ma potrzeby korzystania z zasad leżących u podstaw metod probabilistycznych. Wystarczającym jest wykazanie na konkretnych przykładach, że dana metoda umożliwia uzyskanie satysfakcjonujących rezultatów [6].

Stąd też w naukach medycznych próba badawcza dobierana jest często w sposób nieprobabilistyczny. Przyjęło się wyróżniać cztery podstawowe metody nieprobabilistycznego doboru próby:

Dobór Oparty na Dostępności Danych

Taka metoda nie podlega kontroli nad reprezentatywnością próby. Może być ona uzasadniona jedynie wtedy, gdy badacz chce uzyskać wyniki (zmierzyć rezultaty swoich działań) odnoszące się wyłącznie do tej grupy. W takim przypadku należy jednak zachować ostrożność w uogólnianiu wniosków z tak uzyskanych danych [7,8].

Dobór Celowy Lub Arbitralny

Niekiedy zasadnym jest dobranie próby na podstawie własnej wiedzy o badanej populacji oraz celach badań. Ten sposób doboru nazywany jest próbą celową (purposive sampling) lub arbitralną (judgemental sampling) [7,8].

Metoda Kuli Śnieżnej

(snowball sampling). Metoda jest stosowana w sytuacji, gdy badaczowi trudno jest odszukać członków jakiejś specyficznej populacji. Najpierw badacz dociera do kilku członków badanej populacji, a następnie prosi ich o wskazanie innych osób spełniających określone kryteria. Choć metoda nie spełnia założeń stojących u podstaw probabilistyki, jednak umożliwia dotarcie do osób kompetentnych, którzy pozwolą najtrafniej opisać badany problem [7,8].

Dobór Kwotowy

Dobór kwotowy, to metoda nielosowego doboru próby badawczej. Struktura osób w próbie (liczebność, odsetek, itp.) ustalana jest w taki sposób, aby zachować proporcje ich rzeczywistego udziału w całej badanej populacji. Dla przykładu, jeżeli struktura badanej populacji liczy 51% kobiet i 49% mężczyzn, to takie proporcje powinny zostać przyjęte w doborze próby. Przy doborze kwotowym najczęściej uwzględnia się płeć, wiek oraz miejsce zamieszkania badanych osób. Próba z doboru kwotowego, chociaż nie jest typową techniką losową, to jednak w szczególnych warunkach może być reprezentatywną dla całej populacji. Wszystko zależy od tego, jak szczegółowymi danymi na temat tej populacji dysponujemy. Należy jednak zachować ostrożność w przekładaniu wyników z takiej próby na populację generalną. Wielu statystyków kwalifikuje kwotowy dobór próby jako reprezentatywność przedmiotową, czyli taką, w której zachowana jest struktura kluczowych cech właściwa dla populacji. W określonych sytuacjach uzyskany stopień reprezentatywności może okazać się jednak wystarczający, a uzyskiwane wyniki za satysfakcjonujące [7,9].

Należy zauważyć, że choć powyższe metody nie spełniają kryteriów stawianych metodom probabilistycznym, to jednak w niektórych sytuacjach uzyskane za ich pomocą wyniki mogą być bliższe prawdzie niż w przypadku metod probabilistycznych. Taki przypadek może mieć miejsce w sytuacji, kiedy grupa celowa jest relatywnie mała w stosunku do całości populacji, a ponadto prowadzone rejestry nie zawierają odpowiednich informacji, np. z powodu tego, że dane takie mogą zaliczać się do wrażliwych. Dotarcie do osób spełniających oczekiwane kryteria może być niemożliwe, a losowanie próby z ogółu populacji może prowadzić do sytuacji, w której w badanej próbie nie znajdą się pożądane osoby lub

będzie ich zbyt mało aby uzyskane wyniki można byłoby uznać za w pełni rzetelne. W takim przypadku dobór celowy bądź celowo-losowy może okazać się lepszy od „klasycznie” losowego [9,10].

Skale Pomiarowe

W badaniach naukowych (o charakterze medycznym) podstawowy materiał, który otrzymuje lekarz, składa się z pomiarów i obserwacji przeprowadzanych na pacjentach. Tego typu zbieranie danych przyjęło się określać pojęciem pomiaru. Na potrzeby takich pomiarów opracowano teorię skal pomiarowych oraz zakres ich wykorzystania.

Pomiarem nazywa się czynność przyporządkowania liczb (bądź symboli, które podlegają zliczaniu) przedmiotom lub zdarzeniom zgodnie z przyjętym zbiorem reguł. Takie przyporządkowanie odbywa się zgodnie ze skalą wykorzystywaną do pomiaru. Skala pomiarowa jest to system symboli kodujących wyniki pomiaru lub uzyskane dane. Pomiary licznych cech i własności powinny być obiektywne precyzyjne i odtwarzalne. To, jakie pomiary i obserwacje się stosuje, zależy od rodzaju badania i istniejących symptomów chorobowych.

Przyjęło się wyróżniać (za Stevenssem) następujące podstawowe skale pomiarowe: nominalną, porządkową, przedziałową i ilorazową. Wszystkie cztery skale określają tzw. poziomy pomiaru, które różnią się pod względem sposobu w jaki przyporządkowuje się je cechom badanych obiektów. Skale te charakteryzują się różną precyzją pomiaru oraz dopuszczalnością stosowanych operacji matematycznych i testów [8,10,11].

Skala nominalna - wyraża wyniki obserwacji obiektów (poprzez obiekt rozumie się także osobę uczestniczącą w badaniu (obserwacji)), które mają charakter danych jakościowych. Dane kodowane są na skali nominalnej, kiedy przyjmują wartości (etykiety), dla których nie istnieje uporządkowanie wynikające z natury danego zjawiska. Przykładowe dane nominalne to: płeć, miejsce zamieszkania, miejsce pracy, przebyte choroby, upodobania, itp. Nawet jeśli dane wyrażane w skali nominalnej są kodowane z użyciem liczb, to są one tylko umownymi identyfikatorami i nie można na nich wykonywać żadnych działań arytmetycznych bądź dokonywać porównań. Dopuszczalne operacje matematyczne to zliczanie przypadków zakwalifikowanych do określonej grupy. Kodowanie wyników za pomocą cyfr nie upoważnia do wykonywania działań matematycznych

na takich danych innych niż sumowanie określonych przypadków. W analizach statystycznych dopuszczalne jest wykonywanie testów statystycznych dla danych nominalnych. Najczęściej stosowanym testem jest test χ^2 niezależności oraz χ^2 zgodności. Siłę związku mierzy się za pomocą współczynników kontyngencji V Cramera oraz C Pearsona. Dopuszczalne operacje matematyczne wykonywane na danych to: zliczanie zdarzeń (liczba relacji równości, różności) [1,8].

Skala porządkowa zawiera wyniki obserwacji obiektów które są uporządkowane w zależności od ich rozmiarów lub znaczenia. Skala nie umożliwia określenia różnicy pomiędzy wyodrębnionymi obiektami, a jedynie ich kolejność. Nie można więc w uzasadniony sposób określić różnicy bądź ilorazu pomiędzy dwiema wartościami. Dopuszczalne operacje matematyczne wykonywane na danych w skali porządkowej to: zliczanie zdarzeń oraz wyznaczanie występujących relacji (równości, różności, większości, mniejszości) – bez określania o ile. Przykładowe dane w skali porządkowej to: wykształcenie, miejsce zamieszkania (np. wieś, małe miasto, duże miasto), wszelkiego rodzaju rankingi (kolejność miejsca), rodzaje ryzyka (np. niskie, średnie, wysokie), itp.

Na danych wyrażonych w skali porządkowej nie można więc wykonywać takich działań jak: mnożenie, dzielenie, obliczanie średniej, odchylenia standardowego, regresji liniowej, itp. Dopuszczalne testy statystyczne to: testy nieparametryczne (np. znaków serii, Kołmogorowa-Smirnowa, itp.) [1,8,10].

Skala przedziałowa - nazywana również interwałową - umożliwia przypisanie znaczenia różnicy pomiędzy wynikami obserwacji obiektów. Wyniki obserwacji są umieszczone w określonym przedziale liczbowym, a odległością między obiektami jest różnica między wynikami obserwacji, wyrażona w jednostkach mierzonej wielkości. Dopuszczalne operacje matematyczne wykonywane na danych to: wszystkie operacje wykonywane na skalach niższych (nominalnej i porządkowej) oraz dodawania i odejmowania, obliczanie średniej arytmetycznej, wariancji, odchylenia standardowego, korelacji Pearsona, regresji liniowej i logistycznej oraz mnożenie i dzielenie. Dopuszczalnymi testami statystycznymi są testy parametryczne (np. F -ilorazu wariancji, t-Studenta; test dla dwóch średnich) [1,8].

Skala ilorazowa (nazywana także stosunkową) - zalicza się do skal ilościowych i jest najdokładniejszą z

wymienionych rodzajów. Umożliwia pomiar wartości zmiennej i dokonywanie na niej dowolnych obliczeń matematycznych, w tym operacji mnożenia i dzielenia. Cechą charakterystyczną skali ilorazowej jest to, że posiada naturalny punkt zerowy, poniżej którego nie występują już żadne obserwacje. Do zmiennych, które można mierzyć za pomocą skali ilorazowej, zalicza się m.in. wiek mierzony w latach, wzrost w jednostkach długości (a także inne cechy biometryczne), dochód liczony w jednostkach monetarnych, temperatura mierzona na skali Kelvina, itp. [1,8,10].

Należy zauważyć, że dopuszczalna jest konwersja skal od skali silniejszej do słabszej, lecz w drugą stronę taki zabieg nie jest realizowalny (sztuczne wzmacnianie skali nie jest dopuszczalne). Nie można bowiem z mniejszej ilości informacji zawartej w danych wyrażonych w skali słabszej, spowodować ich zwiększenia, czyli np. z danych nominalnych, lub porządkowych, typu niski średni, wysoki nie można obliczyć różnicy we wzroście osób, które uczestniczyły w badaniu. Można natomiast dane ze skali ilorazowej przekształcić do dowolnej skali słabszej (np. nominalnej), poprzez przyjęcie granic przedziałów i przyporządkowanie do nich poszczególnych obiektów (osób) [1,12,13].

Im mocniejsza jest skala w której dane są kodowane, tym więcej wniosków można z takich badań wyciągnąć, a ponadto są one lepiej udokumentowane. Projektując badania należy więc zwrócić szczególną uwagę na możliwość uzyskania danych, które można wyrazić w możliwie najmocniejszej skali. Niejednokrotnie badacze na etapie analizy uzyskanych rezultatów dochodzili do wniosków, że przyjęcie innego podejścia i odpowiednie przygotowanie badań pozwoliłoby uzyskać więcej przydatnych informacji, rozszerzających spektrum możliwych do zastosowania narzędzi statystycznych. Niestety bardzo często jest już wtedy za późno na zmianę lub uzupełnienie wyników [1,10,13].

WNIOSKI

Już na etapie przygotowania badań należy:

- określić charakter badań,
- przyjęć sposób doboru próby badawczej,

- dobrać skalę w jakich dane będą kodowane,
- wybrać metody analizy uzyskanych rezultatów.

Brak prawidłowo przygotowania badań, może zniweczyć trud ich przeprowadzenia. Im lepiej przygotowane są więc badania, tym większa jest szansa na rzetelne wnioski z nich wypływające.

PIŚMIENNICTWO

1. Babbie E. Badania społeczne w praktyce. Wydawnictwo Naukowe PWN. Warszawa 2003.
2. Brzeziński J. Metodologia badań psychologicznych. Wydawnictwo Naukowe PWN. Warszawa. 1996.
3. Jaworski M. Relationship between health behaviors and quality of life on the one hand and satisfaction with health condition on the other hand in patients with psoriasis. *Our Dermatol Online*. 2013;4:453-7.
4. Stanisław A. Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom 1. Statystyki podstawowe. StatSoft. Kraków 2006.
5. Kovalchuk M, Shkilna M, Andreychyn M, Vasylieva N, Demyanenko V. Optimization of lamblia microscopical diagnostics by the method of polarized fluorescence for patients with rosacea and urticaria. *Our Dermatol Online*. 2013;4:287-9.
6. Ziajka E, Błaszczyk J. Precision of doctor's diagnoses of cancers, pre-cancer conditions atypical and dysplastic lesions. *Our Dermatol Online*. 2018;9(e):e3.1-e3.6.
7. Steczkowski J. Metoda reprezentacyjna w badaniach zjawisk ekonomiczno-społecznych. Wydawnictwo Naukowe PWN. Warszawa-Kraków. 1995.
8. Walesiak M. Metody analizy danych marketingowych. Wydawnictwo Naukowe PWN. Warszawa 1996.
9. Stein ML, Buskens V, van der Heijden PGM, van Steenberghe JE, Wong A, Bootsma MCJ, et al. A stochastic simulation model to study respondent-driven recruitment. *PLoS One*. 2018;13:e0207507.
10. de Vries S, Schaapveld M, Kardaun JW, de Bruin KH, Krol AD, Lugtenburg PJ, et al. Comparing causes of death of Hodgkin lymphoma and breast cancer patients between medical records and cause-of-death statistics. *Clin Epidemiol*. 2018;10:1523-31.
11. Patil V, Joshi A, Noronha V, Bhattacharjee A, Dhumal S, Chandrakanth MV, et al. Quality of life and quality-adjusted time without toxicity in palliatively treated head-and-neck cancer patients. *South Asian J Cancer*. 2018;7:249-53.
12. Jaworski M. Is sexology necessary in dermatology? Review of most common sexual dysfunction in skin disorders. *Our Dermatol Online*. 2017;8:341-7.
13. Jaworski M. Doctors' Support - An important part of medical therapy and quality of life. *Our Dermatol Online*. 2015;6:270-5.

Copyright by Wojciech Jasek, et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Source of Support: Nil, Conflict of Interest: None declared.