

EPINARIUM

NIEFORMALNY BIULETYN
POLSKIEGO STOWARZYSZENIA EPIDEMIOLOGICZNEGO
Numer 6 Grudzień 2010

* * *

*Wszystkim Członkom Polskiego Stowarzyszenia Epidemiologicznego,
jego Sympatykom oraz wszystkim Czytelnikom redakcja Epinarium
życzy Wesółych Świąt i Szczęśliwego Nowego Roku*

* * *

W numerze: **Od Redaktora, Epiforum** (Konferencja Polskiego Stowarzyszenia Epidemiologicznego), **Epiprofil** (Katedra i Zakład Medycyny i Epidemiologii Środowiskowej, Śląski Uniwersytet Medyczny - Zabrze), **Epikon** (Słowniczek Bazy Danych, Kwestionariusz)

OD REDAKTORA

Niniejsze wydanie Epinarium zbiega się z końcem roku 2010. Jest to rok krytyczny dla środowiska naukowego w Polsce i dotyczy to zarówno instytutów naukowo-badawczych, jak i szkół wyższych. Wdrożone już lub planowane reformy mają jednakże szerszy wymiar. Dla naszego środowiska istotne jest, że zapowiadają one także reformę systemu ochrony zdrowia oraz wkraczają w obszar kształcenia na poziomie uniwersyteckim. Proces trwa, a co z niego wyniknie powinniśmy wiedzieć w 2011 roku.

W mijającym roku, po dużych wydarzeniach angażujących nasze środowisko w roku poprzednim, zorganizowaliśmy doroczną Konferencję PSE. Krótka relacja ze spotkania znajduje się w sekcji „Epiforum”. Z kolei w sekcji Epiprofil przedstawiona jest sylwetka Katedry i Zakładu Medycyny i Epidemiologii Środowiskowej z Zabrze. Prezentacja ośrodka obejmuje okres od 1948 roku, a lektura tekstu daje możliwość prześledzenia procesu kształtowania się akademickiego ośrodka epidemiologicznego.

W poprzednim numerze Epinarium znalazło się „Dziesięć przykazań” Daniela W. Byrne’a, biostatystyka z Uniwersytetu Vanderbilt w Nashville (USA). W niniejszym wydaniu pojawia się esej, który omawia zasady przygotowania bazy danych, odwołując się do kluczowych zaleceń Dr. Byrne’a. Autorką eseju jest Theresa A. Scott, bezpośrednia

współpracowniczka Dr. Byrne’a. Jej uwagi mają duże praktyczne znaczenie, a formułowane wskazówki oraz zalecenia są skądinąd znane i stosowane w wiodących ośrodkach epidemiologicznych. Esey T. Scott jest tym bardziej godny uwagi, że pochodzi od osoby posiadającej duże, codzienne i praktyczne doświadczenie w dziedzinie stanowiącej przedmiot zaleceń. W Epinarium zamieszczona jest oryginalna wersja eseju nadesłana przez Autorkę oraz tłumaczenie tekstu na język polski. T. Scott jest statystykiem i dla epidemiologa, zwłaszcza nie prowadzącego analiz statystycznych, istotne może być poznanie poglądów lub wręcz oczekiwań partnerów reprezentujących statystykę. To na pewno poprawi jakość współpracy. Dzisiejszy esej T. Scott jest pierwszą częścią doniesienia jej autorstwa. Jesteśmy umówieni na kontynuację i część II pojawi się w kolejnym wydaniu Epinarium.

Nie ma dobrej epidemiologii bez dobrego kwestionariusza. Truizm, ale jakże aktualny w dobie rozkwitu badań prowadzonych w ramach prac licencjackich, magisterskich i doktorskich na kierunkach kształcenia związanych z ochroną zdrowia. Jakże często spotyka się w tych raportach sformułowanie, że w badaniu wykorzystano „autorski kwestionariusz”. Temu sformułowaniu rzadko towarzyszy refleksja na temat wartości użytego narzędzia. Bywa też, że czytamy o „autorskim kwestionariuszu ankiety”. To ostatnie sformułowanie jest trudną do wytłumaczenia manierą: kwestionariusz jest ankietą, ankietą jest kwestionariuszem. Każdy kwestionariusz jest

kwestionariuszem autorskim, ale nie każdy kwestionariusz autorski jest jednocześnie kwestionariuszem zwalidowanym. Ten ostatni jest kwestionariuszem, który przeszedł naukowy proces oceny i znana jest wartość tego narzędzia badawczego. Warto zatem prześledzić kryteria oceny kwestionariusza i tego dotyczy drugi esej zamieszczony w sekcji „Epikon”.

(Redaktor)

EPIFORUM

Konferencja Polskiego Stowarzyszenia Epidemiologicznego

Konferencja odbyła się w dniu 20 października 2010 r., w Narodowym Instytucie Zdrowia Publicznego – Państwowym Zakładzie Higieny. Program obejmował cztery punkty. Pani doc. dr hab. med. Magdalena Rutkowska z Instytutu Matki i Dziecka w Warszawie wystąpiła z referatem „Długofalowa ocena rozwoju noworodków skrajnie niedojrzałych – badanie własne (PREMATURITAS) na tle badań światowych”, a dr Michał Skrzypek z Wydziału Zdrowia Publicznego Śląskiego Uniwersytetu Medycznego przedstawił „Skutki zdrowotne narażenia na zanieczyszczenia powietrza pochodzenia komunikacyjnego – możliwości dokumentowania ryzyka zdrowotnego”. Dwa różne tematycznie doniesienia łączył wysoki poziom oprawy metodologicznej, a możliwość śledzenia przedstawianych działań i ich wyników była bardzo dobrą okazją do zapoznania się z aktualnie preferowanymi i nowoczesnymi metodami badań epidemiologicznych oraz interpretacją ich wyników. W opinii uczestników oba referaty reprezentowały wysoki poziom, co zrekompensowało fakt, że na udział w Konferencji zdecydowało się tak mało osób. W kolejnym punkcie programu dr Katarzyna Szamotulska przedstawiła sprawozdanie z przebiegu konferencji „EUROEPI 2009”. O tym wydarzeniu Epinarium informowało w jednym z poprzednich numerów. Warto podkreślić, że w warstwie organizacyjnej, podobnie jak w warstwie merytorycznej, „EUROEPI 2009” zakończyło się pełnym sukcesem. Konferencja PSE była także okazją do dyskusji nad przyszłością naszego Stowarzyszenia. Omówiono między innymi możliwość przygotowania, przez nasze środowisko, stanowiska w sprawie populacyjnych programów badań przesiewowych, napotykających – jak wykazują aktualne doświadczenia – szereg trudności realizacyjnych. Inna ciekawa propozycja dotyczyła przygotowania warsztatów poświęconych matematycznemu (statystycznemu) modelowaniu procesów epidemicznych. Osobnym, ale zawsze aktualnym tematem, była kwestia statutowych działań PSE (m.in. lista członków, składki). Informacje na ten temat trafią do Członków PSE.

(JEZ)

EPIPROFIL

Katedra i Zakład Medycyny i Epidemiologii Środowiskowej – Śląski Uniwersytet Medyczny

Historia Katedry jest tak długa, jak długa jest historia kształcenia studentów medycyny na Górnym Śląsku, w powojennej Polsce. W tym czasie Katedra przechodziła różne fazy organizacyjne, ale zawsze funkcjonowała w obszarze higieny i epidemiologii, z aktywnością także w pokrewnych obszarach nauk medycznych. Dzisiaj jednostka jest znana jako Katedra i Zakład Medycyny i Epidemiologii Środowiskowej w Zabrze (Śląski Uniwersytet Medyczny w Katowicach), a jej kierownikiem jest prof. dr hab. n. med. Jadwiga Joško. Katedra jest zlokalizowana w Zabrzu (ul. H. Jordana 19, 41-808 Zabrze).



Katedra powstała w 1948 roku jako **Katedra i Zakład Higieny Ogólnej**, a jej pierwszym kierownikiem był prof. dr med. Brunon Nowakowski, pierwszy rektor Akademii Lekarskiej. Prof. Nowakowski był specjalistą z zakresu higieny pracy, co pierwszorzędnie wpisywało się w zadania stawiane Akademii przez władze wojewódzkie (zapobieganie szkodliwemu oddziaływaniu warunków pracy na zdrowie pracowników, szczególnie w przemyśle górniczym i hutniczym).

W początkowym okresie działalność Katedry koncentrowała się na sprawach organizacyjnych (nabór kadry naukowo-dydaktycznej i pomocniczej, przystosowanie pomieszczeń dydaktycznych, pozyskanie aparatury itp.) Praca dydaktyczna w Katedrze rozpoczęła się w roku akademickim 1949/50 na II roku stomatologii. Później, zgodnie z programami ministerialnymi, zajęcia z higieny dla stomatologii były prowadzone na III roku, a następnie na IV roku. Dla studentów medycyny zajęcia prowadzono na IV roku studiów. Pierwszymi asystentami zostali studenci II roku medycyny Romuald Dukat i Antonina Demczuk. W roku 1950 zespół pracowników Katedry uzupełnił starszy asystent Zakładu Higieny Akademii Medycznej w Krakowie dr Klaudiusz Ciubra, obejmując stanowisko adiunkta oraz asystent Zakładu Higieny Akademii Medycznej w Gdańsku dr Marek Zacharewicz, który został starszym asystentem, a następnie adiunktem. W kolejnych latach stanowiska asystentów obejmowali już absolwenci macierzystej uczelni. W 1958 roku kadre naukowo-dydaktyczną stanowili prof. dr med. Brunon Nowakowski - kierownik Katedry, lekarz Kazimierz Błach - asystent, dr med. Klaudiusz Ciubra - docent, lekarz Romuald Dukat - st. asystent, lekarz Mieczysław Makuch - asystent, lekarz Ewa Nowakowska - asystent, lekarz Edmund Rogala -

asystent, dr med. Marek Zacharewicz - adiunkt,
lekarz Benon Zieleźnik - st. asystent.

Zgodnie z wytycznymi Ministerstwa Zdrowia, pracę naukową Katedry ukierunkowano na zagadnienia związane z higieną pracy. Od 1950 roku główna tematyka badań naukowych dotyczyła mikroklimatu w zakładach pracy, szczególnie zapylenia powietrza w kopalniach węgla kamiennego, oddziaływania promieniowania podczerwonego na pracowników fabryk materiałów ogniotrwałych i hut oraz bakteryjnego zanieczyszczenia powietrza w pomieszczeniach zakładów służby zdrowia. Drugi kierunek badań naukowych wiązał się z epidemiologią gruźlicy jako zagadnieniem higieny społecznej, a trzeci z higieną żywienia i rozwojem fizycznym młodzieży szkolnej.



W 1960 roku, z chwilą przejścia prof. Brunona Nowakowskiego na emeryturę, pełniącym obowiązki kierownika, a od 1963 roku kierownikiem Katedry został doc. dr med. Klaudiusz Ciubra (prof. nadzwyczajny Śląskiej Akademii Medycznej od 1965 r.).

Zajęcia dydaktyczne z przedmiotu „Higiena” w tym okresie prowadzone były dla studentów IV roku studiów Wydziału Lekarskiego i IV roku Oddziału Stomatologicznego. Praca naukowa koncentrowała się w głównej mierze na problemach higieny pracy.

W 1971 roku, w związku ze zmianami strukturalnymi Uczelni, Katedra Higieny została rozwiązana a w jej miejsce powstał **Instytut Medycyny Społecznej**, którego dyrektorem został prof. Ciubra (do uzyskania emerytury w 1979 r.). Instytut składał się z czterech zakładów: Zakład Higieny - kierownik prof. dr med. Klaudiusz Ciubra, Zakład Epidemiologii i Statystyki - kierownik dr n. med. Romuald Dukat, Zakład Dydaktyki Medycznej - kierownik doc. dr med. Benon Zieleźnik i Zakład Organizacji Ochrony Zdrowia - doc. dr n. med. Witold Łobzowski. W 1980 r. dyrektorem Instytutu został doc. Benon Zieleźnik, który pełnił tę funkcję do czasu rozwiązania Instytutu w 1981 roku. Nadal funkcjonował Zakład Higieny w Zabrze ucząc studentów IV roku medycyny i IV roku stomatologii. Praca naukowa prowadzona w Katedrze dotyczyła epidemiologii przewlekłych chorób układu oddechowego, szkodliwego działania związków chemicznych, metali ciężkich i innych związków i czynników w warunkach narażenia zawodowego. Ponadto prowadzono szkolenia dla pracowników zakładów pracy z zakresu toksykologii przemysłowej i higieny pracy.



W 1981 roku w Śląskiej Akademii Medycznej nastąpił powrót do struktury katedralnej. Powołano **Katedrę Higieny i Epidemiologii** z trzema jednostkami: Zakład Higieny, Zakład Epidemiologii i Statystyki oraz Zakład Organizacji i Ochrony Zdrowia. Kierownikiem Katedry i równocześnie Zakładu Higieny został doc. dr n. med. Benon Zieleźnik. W Zakładzie prowadzono zajęcia z „Higieny i Epidemiologii”, a badania naukowe dotyczyły oceny toksyczności nowych związków stosowanych w przemyśle oraz oceny warunków środowiska pracy.

W 1986 roku zlikwidowano Katedrę Higieny i Epidemiologii, a Zakład Epidemiologii i Statystyki włączono w struktury powstałego w 1975 roku Wydziału Lekarskiego w Katowicach przekształcając go w Zakład Higieny i Epidemiologii. Tak więc od 1986 roku funkcjonowały **dwa Zakłady Higieny i Epidemiologii**, jeden w ramach Wydziału Lekarskiego z Oddziałem Stomatologicznym w Zabrze, którego kierownikiem był doc. Zieleźnik, a drugi na Wydziale Lekarskim w Katowicach, którym kierował doc. dr n. med. Romuald Dukat.

Po śmierci doc. Zieleźnika w 1988 roku funkcję kuratora Zakładu Higieny i Epidemiologii w Zabrze sprawował doc. dr n. med. R. Dukat. W następnym roku nowym kierownikiem **Zakładu Higieny i Epidemiologii** został doc. dr hab. n. med. Zygmunt Stęplewski, mianowany następnie na stanowisko



profesora nadzwyczajnego Śląskiej Akademii Medycznej. W 1991 roku Zakład Higieny i Epidemiologii został rozwiązany a w jego miejsce powołano **Katedrę i Zakład Medycyny Środowiskowej i Zapobiegawczej**. Kierownikiem pozostał prof. Stęplewski, który w 1998 roku otrzymał tytuł naukowy. Katedra kontynuowała nauczanie „Higieny i Epidemiologii”, a w późniejszym okresie, również przedmiotu „Zdrowie Publiczne” (VI rok medycyny i III rok stomatologii). W pracach naukowych koncentrowano się na zagadnieniach dotyczących szkodliwego działania czynników fizycznych, chemicznych i psychospołecznych na organizmy zwierzęce i organizm człowieka. Zajmowano się m.in. wpływem stresu, hipertermii i węglowodorów aromatycznych na organizmy zwierzęce i ludzkie oraz ochronnym działaniem

nienasyconych kwasów tłuszczowych na błonę śluzową żołądka w czasie stresu. W kręgu badań był także wpływ ołowiu na organizm dziecka oraz środowiskowe uwarunkowania astmy oskrzelowej i niskiej urodzeniowej masy ciała u śląskich dzieci.

W 1999 roku rozwiązano Katedrę i Zakład Medycyny Środowiskowej i Zapobiegawczej i powołano **Katedrę Medycyny Środowiskowej** składającą się z dwóch zakładów: Zakładu Medycyny i Epidemiologii Środowiskowej oraz Zakładu Zdrowia Publicznego. Kierownikiem Katedry i równocześnie Zakładu Medycyny i Epidemiologii Środowiskowej został prof. dr hab. n. med. Zygmunt Stęplewski a kierownikiem Zakładu Zdrowia Publicznego dr hab. n. med. Antoni Węgiel. Kontynuowano tematykę badawczą z poprzedniego okresu a zajęcia dydaktyczne prowadzono w oparciu o programy dla poszczególnych kierunków studiów. Po kolejnej reorganizacji w 2002 roku ze struktury Katedry Medycyny Środowiskowej wyłączono Zakład Zdrowia Publicznego, który wszedł w struktury organizacyjne nowego Wydziału Zdrowia Publicznego w Bytomiu.



W nowej sytuacji podjęto decyzję o rozwiązaniu Katedry i powołaniu, w jej miejsce, **Katedry i Zakładu Medycyny i Epidemiologii Środowiskowej**. Pierwszym kierownikiem Katedry i Zakładu został prof. dr hab. n. med. Zygmunt Stęplewski, który 1 października 2002 roku przeszedł na emeryturę, a nowym kierownikiem

została dr hab. n. med. Jadwiga Joško. W maju 2010 roku powołano ją na stanowisko profesora nadzwyczajnego Śląskiego Uniwersytetu Medycznego, a w październiku tego roku otrzymała tytuł naukowy profesora nauk medycznych.

Kontynuując dotychczasową działalność naukowo-badawczą Katedry rozszerzono zainteresowania m.in. o takie zagadnienia jak: epidemiologia narkomanii, alkoholizmu, nikotynizmu, bezsenności, otyłości wśród dzieci i młodzieży oraz o problematykę angiogenezy w przewodzie pokarmowym i ośrodkowym układzie nerwowym w stresie kontrolowanym i niekontrolowanym, w ostrym stresie, po krwotoku podpajęczynówkowym oraz w regeneracji ośrodkowego układu nerwowego. Obecnie w Katedrze pracuje 1 osoba z tytułem profesora, 4 osoby z tytułem doktora, 2 starszych techników, 1 technik i dwie osoby obsługi na ½ etatu.

(Materiał opracowany przez Zespół Katedry: Mirosław Tyrpień, Jadwiga Joško, Janusz Kasperczyk, Beata Malara)

EPIKON

An Elaboration of the Guidelines for Data Collection & Data Entry: Part I – Your Data Dictionary

The goal of this article is to elaborate on Dan Byrne's "10 Data Entry Commandments" that were presented in that last issue as well as provide additional considerations when collecting and entering data.

To me, the procedure of data collection and data entry includes two major steps: (1) creating your data dictionary; and (2) creating your data file(s). Let's first discuss the creation of your data dictionary, which is a document that includes a description of the study variables as well as data management procedures. A data dictionary is most advantageous if defined *before* any data is collected.

When defining your data dictionary, it is beneficial to write a detailed list of the information to be collected and the concepts to be measured in the research study. This list should directly relate to the specific aim(s) of the *current* study.¹ You also want to ensure that the list includes all the information needed to (1) describe the *sample* of "subjects" you will study; and (2) perform the planned statistical analysis. Thinking of the following situation often helps the researcher – if I already had all my data in hand and was writing the manuscript, what results would I want to present and what would I want to be able to conclude from the results? Thinking to the future often clarifies the necessary items to collect, the key comparisons to be made, and the necessary statistical analyses to be performed.

This detailed list of items (ie, variables) is then converted to the data dictionary. For each variable, the data dictionary includes the:

- variable name,
- role of the variable (in the planned statistical analysis),
- variable label,
- unit of measurement (if applicable),
- type of variable, and
- permissible values or range of values.

The variable name is used to identify the variable in the data file(s) and must be unique. It should be short but understandable and self-explanatory. As Dan hinted to, many statistical software programs

¹ Many researchers have the habit of collecting extraneous measures, which they hope will lead to future research studies, but most often are just an additional fruitless burden for the current research study. The researcher will have much more success if they collect the extraneous measures in a separate appropriately designed research study.

have variable name restrictions – most often, variable names may consist only of letters (A-Z and a-z), digits (0-9), periods (.), and the underscore (_). Special characters such as +, -, *, /, and ? are not allowed in variable names. And DON'T FORGET, many statistical software programs are case sensitive so 'age' and 'Age' are two distinct variable names.

The role of each variable can be 'outcome', 'predictor', 'confounder', or 'additional descriptor' – information that dictates the statistical analysis to be performed. The outcome is the variable that is the focus of the research study, whose variation or occurrence you are seeking to understand. The predictor is the variable that may influence the size or the occurrence of the outcome (exposure variable or risk factor). The researcher is most often interested in the possible association/relation that exists between the outcome and the predictor. We also must consider any confounders – variables that differ between values of the predictor variable, which also affect the outcome. Both the predictor-confound and confounder-outcome associations are needed to correctly estimate the predictor-outcome association. Most research studies have one or more outcomes, predictors, and confounders. Any 'additional descriptor' is used to describe the sample of 'subjects' you studied.

The variable label is the 'pretty' label used to fully describe the variable. For example, 'age0' may be the variable name, but 'Age at baseline (years)' is the variable label. While the variable names will be used in performing the analysis in the statistical software program, the variable labels allow the statistician to better understand what each variable represents. The statistician can also incorporate the variable labels into the analysis output, making the results easier to interpret. As shown above, the variable label often contains the unit of measurement, if applicable. In addition to adding clarity to the results, clarifying the units of measurement also allows the statistician to check for possible data entry errors.

Like the role of the variable, the type of variable also dictates the statistical analysis to be performed. A continuous variable is one that has any number of possible values (eg, weight). A variable can also be discrete numeric, where the set of possible values is a finite (ordered) sequence of numbers (most often integer values; eg, a pain scale of 0, 1, 2, ..., 10). A categorical variable is one that has only certain possible values, which are most often not numeric (eg, race as 'Caucasian', 'African American', 'Asian', or 'Other'). A binary (dichotomous) variable is a categorical variable with only two possible values (eg, gender as 'Male' or 'Female'). Lastly, an ordinal variable is a categorical variable for which there is a definite ordering of the categories (eg, severity of

lower back pain as 'none', 'mild', 'moderate', and 'severe').

As a statistician, it is also important to know the permissible values or range of values for each variable. The permissible values define the values a categorical variable can take. Permissible values can be coded as numeric or text values – for example, gender (a binary variable) could be coded as 0 and 1 for 'Female' and 'Male' or as 'F' and 'M'. Most statistical software programs allow either numeric or text coding of categorical variables, while others (such as SPSS) require all categorical variables to be numerically coded. For continuous and discrete numeric variables, we define the permissible range of values (the allowable minimum and maximum values), which is used to guide data editing – values outside the defined range are flagged as possible data errors and checked for accuracy.

The data dictionary should also include definitions of redefined or derived variables. For example, body-mass-index (BMI) derived from height and weight or disease status derived from ten clinical variables. Most statisticians will re-calculate any derived variables in the data files to ensure their accuracy. With this in mind, researchers will often collect only the variables needed to calculate any derived variables, allowing the statistician to perform the (sometimes very complicated) calculations using a statistical software program.

Lastly, the data dictionary should include any branching logic that will be used for specific variables. For example, the researcher plans to collect 'type of diabetes' using two variables. First, the subject will be asked 'Do you have diabetes?' (permissible values: 'No' and 'Yes'). If the subject says 'Yes', then they will be asked 'Do you have Type I or Type II diabetes?' (permissible values: 'Type I' and 'Type II'). In this example, a subject who responds 'No' to 'Do you have diabetes?' will have a missing value for their second 'type of diabetes' variable. Branching logic is used to thread through a series of variables based on the values of each variable – as in the example, the 'subjects' for which the thread does not apply will have missing values for specific variables in the thread. Documenting this branching logic is extremely helpful for the statistician receiving your data file(s). As with other

information contained in the data dictionary, the statistician can double-check all of the branching logic used to collect the data, flagging possible data errors and inconsistencies.

The following is a portion of an example data dictionary:

Name	Role	Label	Units	Type	Values
group	Predictor	Treatment		Binary	1 = Placebo; 2 = Treatment
age	Predictor	Age	Years	Continuous	18 - 75
sex	Predictor	Gender		Binary	1 = Female; 2 = Male
ht	Predictor	Height	inches (in)	Continuous	48 - 96
wt	Predictor	Weight	pounds (lbs)	Continuous	75 - 350
hct	Predictor	Heart rate	Beats/min	Continuous	30 - 50
bpsys	Predictor	Systolic BP	mmHg	Continuous	100 - 160
bpdias	Predictor	Diastolic BP	mmHg	Continuous	80 - 150
stage	Predictor	Stage of cancer		Discrete numeric	1 - 4
race	Predictor	Race		Categorical	1 = White; 2 = Black; 3 = Other
date1	Additional	Date of surgery			mm/dd/yyyy
complic	Outcome	Complications?		Binary	0 = No; 1 = Yes

There are a few additional considerations worth mentioning. As Dan indicated in his commandments, when collecting and entering continuous variables, it is best to keep them continuous (ie, don't categorize). When continuous variables are collected already categorized (eg, age collected as ≤ 30 or > 30), the original continuous values cannot be recovered and it is nearly impossible to recode with new categories (eg, recode age as ≤ 24 or > 24).

It is also worthwhile to consider how missing data will be recorded. Most statistical software programs use a single value to represent missing values (eg, 'NA', '.', or a blank cell). Possible problems may arise, however, if you choose to represent missing values with a specific numeric value. For instance, if missing data is coded with a value of '99' and the statistician is not aware of this, a subject who has a missing age value may be analyzed as if their age was 99 years. To avoid any potential problems, it is advisable to specify how missing values are represented in the data dictionary.

Last but not least, variables with non-mutually exclusive values are a nightmare to analyze if not defined in a particular way. An example of a variable with non-mutually exclusive values is post-operative maternal complications following a Cesarean section (C-section). The complications of bleeding, high blood pressure, and fever, for instance, can occur in any combination. Most researchers would collect this variable as a text field, with all possible iterations as its values - eg, 'bleeding', 'high BP and fever', 'fever and blood loss', 'BP > 120; bled all night', etc. It can take several hours for the statistician to pull these free text combinations apart and make the subsequent derived variables consistent. A better approach to variables with non-mutually exclusive values is to code them as separate Yes/No binary categorical variables - one for each non-mutually exclusive value. So, for our maternal complications example, the data dictionary would code three separate

Yes/No variables of bleeding, high blood pressure, and fever.

In closing, the data dictionary should be revised as the details of the research study are refined. Also, as Dan mentioned, have a statistician review the data dictionary before any data is collected. Finally, get in the habit of sending the statistician the data dictionary in addition to the raw data files once all of the data is collected - this will prevent an onslaught of clarifying questions regarding your data.

As the title of this article suggests, there will be a follow-up article that will discuss the creation of the data file(s) for a research study.

(Theresa A. Scott, MS, Department of Biostatistics, Vanderbilt University)

Omówienie Zasad Gromadzenia i Wprowadzania Danych. Część I - Słowniczek Bazy Danych *(tekst jest nieautoryzowanym tłumaczeniem powyżej zamieszczonego eseju - polskie tłumaczenie nie obejmuje przykładu słowniczka zamieszczonego w oryginalnym tekście)*

Celem artykułu jest omówienie „10 Przykazań Wprowadzania Danych” Dan’a Byrne’a, przedstawionych w ostatnim numerze, a także przekazanie dodatkowych uwag, które mają znaczenie w fazie gromadzenia i wprowadzania danych.

W procesie gromadzenia i wprowadzania danych wyróżniam dwa ważne działania: (1) przygotowanie własnego słowniczka bazy danych i (2) przygotowanie własnego pliku z danymi. Warto najpierw omówić przygotowanie własnego słowniczka bazy danych, który jest dokumentem zawierającym zarówno opis zmiennych stosowanych w badaniu jak i procedur stosowanych przy manipulacji zmiennymi. Słowniczek bazy danych jest

najbardziej przydatny, gdy został przygotowany przed rozpoczęciem gromadzenia danych.

W fazie przygotowania słowniczka korzystne jest stworzenie szczegółowej listy informacji, które będą zbierane oraz zagadnień, które będą poddane badaniu naukowemu. Lista taka powinna bezpośrednio odnosić się do specyficznego celu(ów) aktualnego badania¹. Warto także upewnić się, że taka lista zawiera informację niezbędną dla (1) opisu próby badanej i (2) dla przeprowadzenia planowanej analizy statystycznej. Przy przygotowaniu takiej listy pomocna może być następująca refleksja – *gdy już zgromadzę wszystkie dane i mam zamiar przystąpić do pisania manuskryptu pracy to jakie wyniki chciałbym zamieścić i jakiego rodzaju wnioski będę w stanie wyciągnąć z wyników mojego badania ?* Myślenie o przyszłych działaniach często ułatwia decyzję w sprawie zakresu zbieranych danych, kluczowych porównań oraz niezbędnych statystycznych analiz.

Szczegółowa lista danych (tj zmiennych) służy do przygotowania słowniczka bazy danych. Dla każdej zmiennej słowniczek zawiera następujące informacje:

- nazwa zmiennej
- rola zmiennej (funkcja w planowanej analizie statystycznej)
- etykieta zmiennej
- jednostka pomiaru (o ile ma zastosowanie)
- typ zmiennej
- dopuszczalne wartości lub zakres wartości zmiennej

Nazwa zmiennej służy identyfikacji zmiennej w bazie danych i musi być unikalna (w tej bazie). Powinna być krótka, ale zrozumiała i informatywna. Dan (Byrne) podkreśla, że wiele statystycznych programów stosuje restrykcje odnośnie dopuszczalnych nazw zmiennych – najczęściej nazwa ma składać się z liter (A-Z i a-z), cyfr (0-9), znaku kropki (.) i podkreślnika (_). Specyficzne znaki, takie jak +, -, *, / i ? nie są dopuszczalne. I pamiętajmy przy tym, że wiele statystycznych programów inaczej traktuje duże i małe litery, co sprawia, że np. nazwy 'wiek' i 'Wiek' są odczytywane jako dwie różne zmienne.

Zmienna może pełnić rolę 'zmiennej krytycznej', 'zmiennej objaśniającej', 'zmiennej zakłócającej' lub 'zmiennej dodatkowej' – rola zmiennych dyktuje dobór procedur statystycznych w ich analizie.

¹ wielu badaczy posiada nawyk pomiaru i gromadzenia dużej liczby zmiennych, z nadzieją na ich wykorzystanie w przyszłości, ale najczęściej stanowią one bezowocne obciążenie aktualnie prowadzonej analizy. O wiele skuteczniejsze jest gromadzenie takich dodatkowych danych podczas osobnego, odpowiednio zaplanowanego badania.

'Zmienna krytyczna' jest tą zmienną, wokół której koncentruje się cel badania i której występowanie lub zmienność planowane badanie ma wyjaśnić. 'Zmienna objaśniająca' jest taką zmienną, której obecność może wpływać na występowanie lub natężenie 'zmiennej krytycznej' (tu: zmienna reprezentująca narażenie lub czynnik ryzyka). Najczęściej w centrum dociekań jest potencjalna zależność/relacja pomiędzy 'zmienną krytyczną' i 'zmienną objaśniającą'. Konieczne jest także uwzględnienie 'zmiennych zakłócających' – zmiennych, które zależą od wartości 'zmiennej objaśniającej' i które także wpływają na 'zmienną krytyczną'. Dla prawidłowego oszacowania zależności 'zmiennej krytycznej' od 'zmiennej objaśniającej' konieczne jest uwzględnienie zarówno relacji 'zmienna objaśniająca' - 'zmienna zakłócająca', jak i relacji 'zmienna zakłócająca' - 'zmienna krytyczna'. W większości badań naukowych występuje jedna lub więcej 'zmiennej krytycznej' i 'zmiennej objaśniającej' oraz jedna lub więcej 'zmiennej zakłócającej'. 'Zmienne dodatkowe' są wykorzystywane przy opisie badanej grupy („charakterystyka badanych”).

Etykieta zmiennej jest pełną nazwą zmiennej, opisującą czytelnie jej znaczenie. Na przykład, 'wiek0' może być nazwą zmiennej, której pełne znaczenie wyjaśnia etykieta 'wiek w momencie rozpoczęcia obserwacji (lata)'. Nazwa zmiennej (kod) jest wykorzystywana w trakcie analizy statystycznej, podczas gdy etykieta zmiennej pozwala statystykowi na lepsze zrozumienie, czym w istocie jest dana zmienna. Statystyk może przy tym wykorzystać etykiety zmiennych przy przygotowaniu raportu prezentującego wyniki analiz, co ułatwi interpretację zamieszczonych w nim wyników. Jak podano w powyższym przykładzie etykieta często zawiera informację o jednostce pomiaru. Poza tym, że zwiększa to czytelność raportu statystycznego taka informacja umożliwia statystykowi sprawdzenie, czy analizowana baza nie zawiera błędnych danych.

Podobnie jak to ma miejsce w przypadku roli zmiennej również jej typ określa rodzaj analiz statystycznych. Zmienna ilościowa jest taką, która posiada jedną z wielu możliwych wartości (np., masa ciała). Zmienna może być dyskretna, gdy jej wartość pochodzi z zamkniętego zestawu uporządkowanych liczb (np. natężenie bólu w skali 0, 1, 2, ..., 10). Zmienna jakościowa posiada tylko niewielkie spektrum możliwych wartości, które najczęściej nie mają charakteru liczbowego (np. rasa kaukaska, afro-amerykańska, azjatycka lub inna). Zmienna binarna (dychotomiczna) przybiera tylko jedną z dwóch wartości (np. płeć męska lub żeńska). W końcu zmienna porządkowa jest zmienną jakościową dla której istnieje hierarchicznie uporządkowany zestaw wartości jakościowych/kategorii (np. natężenie bólu krzyża: brak, małe, średnie, duże).

Dla statystyka ważna jest informacja, jakie są dopuszczalne wartości i jaki jest zakres wartości każdej z analizowanych zmiennych. Dopuszczalne wartości odnoszą się do zmiennych jakościowych. Mogą być one kodowane zapisem numerycznym lub tekstowym – np. płeć (zmienna binarna) może być kodowana jako '0' i '1' dla kobiet i mężczyzn lub jako 'K' i 'M'. Większość programów statystycznych pozwala na zastosowanie obu rodzajów kodowania, podczas gdy niektóre (np. SPSS) wymagają numerycznego formatu dla zmiennych jakościowych. W przypadku zmiennych ilościowych i dyskretnych numerycznych definiuje się dopuszczalny zakres wartości (dopuszczalne minimum i maksimum) – wartości wykraczające poza ten zakres będą oznaczane jako możliwe błędy na etapie wprowadzania danych i wymagają sprawdzenia.

Słowniczek bazy danych powinien także zawierać definicję zmiennych transformowanych lub pochodnych. Przykładem może być wskaźnik masy ciała (BMI) obliczony na podstawie masy i wysokości ciała lub stan zdrowia definiowany przez zmienne reprezentujące wyniki badań klinicznych. Większość statystyków – w trosce o jakość bazy danych – odtworzy procedury, które stosowano dla pozyskania transformowanych zmiennych. W związku z tym badacze powinni ograniczyć się do gromadzenia wyłącznie tych zmiennych źródłowych, które są niezbędne dla pozyskania zmiennych pochodnych, przekazując tym samym statystykom zadanie wykonania (niekiedy bardzo skomplikowanych) obliczeń z wykorzystaniem programu statystycznego.

Wreszcie, słowniczek bazy danych powinien zawierać informację o logicznych procedurach stosowanych w celu pozyskania specyficznych zmiennych. Na przykład, w odniesieniu do cukrzycy można zaplanować – w kwestionariuszu – zestaw dwóch pytań. Pierwsze pytanie brzmi „czy badany choruje a cukrzycę?” (dopuszczalne wartości zmiennej: 'Nie' i 'Tak'). W przypadku odpowiedzi 'Tak' drugie pytanie brzmi „czy to jest cukrzyca typu I czy typu II?” (dopuszczalne wartości zmiennej: 'Typ I' i 'Typ II'). Gdy badany odpowiada 'Nie' na pierwsze z pytań w przypadku zmiennej odzwierciedlającej drugie pytanie pojawi się brak danych. Jednoznaczny opis tej sytuacji przy pomocy logicznej procedury jest niezmiernie przydatny w pracy statystyka – w ten sposób uniknie on błędnej klasyfikacji logicznego braku danych jako braku danych z innej przyczyny. Podobnie, jak to ma miejsce w przypadku innych informacji zawartych w bazie danych, statystyk dysponujący opisem logicznych procedur będzie mógł zweryfikować poprawność danych i zaznaczyć możliwe błędy oraz niekonsekwencje.

Warto także zwrócić uwagę na szereg dodatkowych kwestii. Zgodnie z tym co Dan (Byrne) zawarł w

swoich „10 Przykazaniach” w przypadku gromadzenia danych ilościowych i wprowadzania ich do bazy danych najlepiej jest utrzymywać ich postać ilościową (tj. bez przekształceń w zmienną jakościową). W przypadku gdy transformację zmiennej ilościowej do zmiennej jakościowej przeprowadzono przed jej wprowadzeniem do bazy danych (np. wiek zdefiniowany jako '=<30' lub '>30') nie ma już możliwości (w fazie analizy danych) powrotu do oryginalnej postaci lub przekodowania w celu testowania nowych kategorii (np. wiek przekodowany w kategorii '=<24' lub '>24').

Warto także rozważyć sposób zapisu brakujących danych. W większości programów statystycznych stosuje się w tym celu pojedyncze znaki (np. 'BD' lub puste pole). Problemy mogą się jednakże pojawiać, gdy zastosowany kod ma postać wartości liczbowej. Na przykład, gdy dla wieku brakująca informacja zostanie oznaczona zapisem '99' niepowiadomiony o tym statystyk może traktować tę wartość jako rzeczywisty wiek badanego. Aby uniknąć potencjalnych problemów należy – w słowniczku bazy danych – ujawnić, jak brakujące dane są reprezentowane w bazie danych.

Na koniec, co nie oznacza że jest to sprawa mniejszej wagi, trzeba się odnieść do zmiennych, które nie posiadają wzajemnie wykluczających się wartości. Analiza takich zmiennych, gdy nie są one zdefiniowane we właściwy sposób, może być koszmarem. Przykładem zmiennej o niewykluczających się wzajemnie wartościach jest zmienna odzwierciedlająca 'komplikacje po cięciu cesarskim'. Dla przykładu, w tym przypadku krwotok, wzrost ciśnienia tętniczego i gorączka mogą wystąpić w jakiegokolwiek kombinacji. Większość badaczy pozyskuje a następnie odnotowuje tę zmienną w postaci zapisu tekstowego, stosując wszystkie możliwe warianty zapisu – np. 'krwotok', 'wysokie ciśnienie i gorączka', 'gorączka i utrata krwi', 'RR>120; krwawienie całonocne', itp. Statystyk może spędzić wiele godzin rozdzielając te dowolne zapisy i przetwarzając je w jednorodne zmienne. Lepszym rozwiązaniem w przypadku zmiennych o niewykluczających się wzajemnie wartościach jest kodowanie ich w sposób niezależny jako kolejne, pojedyncze zmienne jakościowe z binarną wartością 'Tak/Nie', dla każdej składowej. W rezultacie, w przypadku przytoczonego przykładu słowniczek bazy danych powinien uwzględniać trzy oddzielne zmienne z wartościami 'Tak/Nie', osobne dla opisu krwotoku, wysokiego ciśnienia tętniczego i gorączki.

W podsumowaniu należy podkreślić, że słowniczek bazy danych powinien być ponownie przejrany gdy uzgodniono i znane są szczegółowe elementy badania naukowego. Zgodnie z tym, co zaleca Dan (Byrne) należy skonsultować słowniczek bazy danych ze statystykiem, zanim przystąpi się do gromadzenia danych. Wreszcie, należy za regułę

postępowania przyjąć zwyczaj przekazywania statystykowi słowniczka bazy danych razem z surowymi danymi, po ich zgromadzeniu. Takie postępowanie pozwoli na uniknięcie serii wielu nieuchronnych prośb o wyjaśnienia i pytań, formułowanych przez współpracującego statystyka.

Zgodnie z tym, co sugeruje tytuł niniejszego doniesienia przewidziana jest kolejna publikacja, w której omówione będą kwestie przygotowania bazy danych.

(Theresa A. Scott MS, Departament of Biostatistics, Vanderbilt University; tłumaczenie JEZ)

Kwestionariusz

Kwestionariusz jest podstawowym narzędziem badawczym w epidemiologii. Z technicznego punktu widzenia jest to proste narzędzie, stanowiące zestaw pytań (niekiedy urozmaiconych prezentacjami obrazkowymi), które mają charakter zamknięty lub otwarty. Pytania zamknięte to takie, którym towarzyszy lista wydrukowanych odpowiedzi (zamknięta lista - stąd nazwa), a badany jest proszony o wskazanie jednej lub więcej niż jednej odpowiedzi najlepiej charakteryzującej jego sytuację, samopoczucie, doświadczenia itp. (*np. czy palisz obecnie papierosy ? tak - nie*). Pytaniom otwartym nie towarzyszy żadna propozycja odpowiedzi - rolą badanego jest udzielenie (wpisanie) właściwej odpowiedzi (*np. jakie obecnie stosujesz leki ? :*). Praca z pytaniami otwartymi może sprawiać trudności badanemu i - może nawet większe - badaczowi. Zróżnicowanie odpowiedzi w tych przypadkach, ich mnogość, a także oddanie inicjatywy badanemu skutkuje dobrze znanymi problemami na etapie analizy danych, nie mówiąc o ich interpretacji. Nie ryzykując wiele można nawet przypuszczać, że informacja zdobyta tą drogą bywa pomijana, co jest usprawiedliwione rozmiarem szumu informacyjnego generowanego przez odpowiedzi na pytania otwarte. W związku z tym pytania tego typu warto stosować oszczędnie i głównie w odniesieniu do kwestii opisywanych przy użyciu ilościowej skali (*np. pytanie o wiek, liczbę lat palenia tytoniu, itp.*).

Kwestionariusz jest narzędziem badawczym i jako narzędzie badawcze powinien zapewnić miarodajny i powtarzalny sposób oceny zjawiska, które bada. Tym zjawiskiem może być opis stylu życia, stanu materialnego, przebytych chorób, aktualnych dolegliwości, itp. Uzyskana informacja może być wykorzystana nawet w konkretnych celach diagnostycznych.

Dobrym przykładem jest tutaj Kwestionariusz Objawów ze Strony Układu Oddechowego, opracowany w 1960 roku pod egidą British Medical Council. Zawiera on szereg specyficznych pytań, wśród których duże znaczenie mają pytania dotyczące objawów kaszlu i odkrztuszania płucociny:

KWESTIONARIUSZ OBJAWÓW ZE STRONY UKŁADU ODDECHOWEGO (fragment)	
1. Czy kaszle Pan(i) zwykle po wstaniu z łóżka zimą lub jesienią ?	TAK / NIE
2. Czy kaszle Pan(i) zwykle w ciągu dnia lub w nocy zimą lub jesienią ?	TAK / NIE
<i>Jeśli odpowiedź TAK na pytanie 1 lub 2 to:</i>	
3. Czy kaszle Pan(i) w ten sposób przez większość dni w tygodniu przez co najmniej 3 miesiące w roku ?	TAK / NIE
<i>Jeśli odpowiedź TAK na pytanie 3 to:</i>	
4. Ile lat kaszle Pan(i) w ten sposób ?	LATA : <input type="checkbox"/>
5. Czy odkrztusza Pan(i) flegmę zwykle po wstaniu z łóżka zimą lub jesienią ?	TAK / NIE
6. Czy odkrztusza Pan(i) flegmę zwykle w ciągu dnia lub w nocy zimą lub jesienią ?	TAK / NIE
<i>Jeśli odpowiedź TAK na pytanie 5 lub 6 to:</i>	
7. Czy odkrztusza Pan(i) flegmę w ten sposób przez większość dni w tygodniu przez co najmniej 3 miesiące w roku ?	TAK / NIE
<i>Jeśli odpowiedź TAK na pytanie 7 to:</i>	
8. Ile lat odkrztusza Pan(i) flegmę w ten sposób ?	LATA : <input type="checkbox"/>

Przytoczony zestaw pytań umożliwia epidemiologiczne rozpoznanie przewlekłego zapalenia oskrzeli, które jest definiowane jako utrzymujący się od co najmniej dwóch lat kaszel z odkrztuszaniem, występujący przez większość dni w tygodniu, przez okres co najmniej trzech miesięcy w roku. Z technicznego punktu widzenia rozpoznanie przewlekłego zapalenia oskrzeli (PZO) stawia się w przypadku, gdy osoba badana udzieli odpowiedzi "TAK" na pytania 3 i 7 oraz co najmniej "2" na pytania 4 i 8 (PZO=TAK, gdy 3=TAK i 7=TAK i 4>=2 i 8>=2). Wybór tych pytań (i odpowiedzi na nie) wynikał ze skrupulatnej, systematycznej i prospektywnej oceny klinicznej osób deklarujących te dolegliwości. Okazało się, że istotne znaczenie kliniczne posiada stan opisany przy użyciu tych, a nie innych (także uwzględnianych i analizowanych) pytań. Przedstawiona procedura ustalania wartości diagnostycznej kwestionariusza jest procedurą standaryzacyjną. W rezultacie epidemiologiczna diagnoza PZO uzyskała rangę diagnozy stosowanej w skali globalnej. Przedstawiony kwestionariusz jest kwestionariuszem standardowym, co oznacza że jego wartość diagnostyczna została zweryfikowana przy użyciu zewnętrznego standardu - w tym przypadku

obserwacji klinicznej ukierunkowanej na dokładną diagnostykę stanu klinicznego, który ma być identyfikowany badaniem kwestionariuszowym.

Standaryzacja kwestionariusza dla celów klinicznych jest procesem żmudnym i czasochłonnym. W badaniach epidemiologicznych tego typu ocena nie zawsze jest konieczna, ale konieczne jest określenie innych atrybutów wartości kwestionariusza. Niezależnie od wielu okoliczności warunkujących wartość badania epidemiologicznego jego ostateczne wyniki są przynajmniej tak dobre, jak dobre jest narzędzie badawcze i tylko, gdy znana jest jego wartość można interpretować wyniki badania, z uwzględnieniem ograniczeń stworzonych przez ograniczone możliwości pomiaru.

O wartości kwestionariusza decyduje wiele parametrów, które w literaturze przedmiotu omawiane są w sposób szczegółowy, często w rozproszonych materiałach. W praktyce najważniejsze znaczenie posiadają dwa atrybuty: akceptowalność i wiarygodność kwestionariusza.

Akceptowalność kwestionariusza

Doświadczenia własne (i nie tylko) uczą, że opinia badacza i opinia osoby badanej na temat kwestionariusza mogą się znacząco różnić. Rozbudowany i wielostronicowy kwestionariusz może zniechęcać badanego do udziału w badaniu lub skłaniać do udzielania pobieżnych odpowiedzi na kolejne pytania. Podobny efekt uboczny będzie konsekwencją zastosowania stylu mało zrozumiałego dla badanego, wykorzystania specjalistycznego słownictwa, niejednoznacznego formułowania pytań, lub - czego nie można pominąć - stawiania pytań uznawanych przez badanego za zbyt osobiste lub nawet intymne. Omawiane zagrożenia dotyczą zwłaszcza kwestionariuszy wypełnianych samodzielnie przez badanych. Potencjalnie dotyczą one zarówno tzw. własnych kwestionariuszy, jak i kwestionariuszy tłumaczonych z języka obcego.

Jedyną metodą umożliwiającą zbadanie akceptowalności kwestionariusza jest poznanie opinii osób postronnych, nie posiadających profesjonalnej, specjalistycznej wiedzy w obszarze stanowiącym zakres tematyczny kwestionariusza. Biorących udział w takim ćwiczeniu trzeba spytać o ich zdanie na temat długości, zrozumiałości tekstu itp., a ich opinię wykorzystać podczas przygotowania finalnej wersji kwestionariusza - oczywiście nie kosztem aspektów merytorycznych narzędzia.

Wiarygodność kwestionariusza

Miarą wiarygodności narzędzia badawczego jest jego zdolność do identyfikacji zjawiska, dla którego to pomiaru zostało ono opracowane i wdrożone. Ocena

wiarygodności (walidacja) jest - jak wynika z lektury podręczników metodologii - zadaniem pracochłonnym i wieloetapowym. Stosując uproszczony schemat takiej oceny można przyjąć, że procedura ta obejmuje ocenę rzetelności, trafności i dokładności:

$$\text{wiarygodność} = \text{dokładność} + \text{rzetelność} + \text{trafność}$$

Dokładność

Kwestia dokładności nie może być jednoznacznie rozstrzygnięta w przypadku badania kwestionariuszowego. Z pewnością pytania proste i zrozumiałe zezwalają na udzielenie bardziej dokładnych odpowiedzi. W wielu sytuacjach nie są to jednakże wystarczające wyznaczniki - trudno przecież o dokładny pomiar subiektywnych odczuć (np. jakość snu, nasilenie duszności, uczucie zmęczenia), ale także kłopoty może sprawiać odpowiedź na pytania dotyczące bardziej mierzalnych zdarzeń (np. regularność wysiłku fizycznego, częstość spożycia ryb, przebyte choroby). Niezależnie od tego czy dany stan jest analizowany przez jedno czy przez kilka pytań w kwestionariuszu dokładność pozostaje trudno mierzalnym stanem.

Rzetelność

Rzetelność jest atrybutem odnoszącym się do powtarzalności wyniku pomiaru danego zjawiska. Narzędzie badawcze jest rzetelne, gdy zastosowane kilkakrotnie do pomiaru konkretnego zjawiska o stabilnym przebiegu w czasie dostarcza, za każdym razem, takiego samego wyniku. W odniesieniu do kwestionariuszy miarą jego rzetelności jest zgodność odpowiedzi na te same pytania, zadane dwukrotnie (lub częściej). Brak zgodności odpowiedzi ma różne przyczyny - może wynikać z nieprecyzyjnego, nieostrego sformułowania pytania, z pośpiechu podczas wypełniania kwestionariusza lub wręcz z nierzetelnych odpowiedzi udzielanych przez osoby badane. Akceptowalny kwestionariusz i czytelna instrukcja jego wypełnienia powinny ograniczyć, ale nie wyeliminują problemu.

W praktyce badanie powtarzalności odpowiedzi na pytania kwestionariuszowe polega na dwukrotnym rozprawdzeniu kwestionariusza u tych samych badanych, w odstępie nie przekraczającym 1-2 tygodni (jest mało prawdopodobne, aby w tym czasie doszło do istotnej zmiany stanu będącego przedmiotem badania). Trudno jest określić wielkość grupy biorącej udział w ocenie powtarzalności, ale nie powinna ona być mniejsza niż 100 osób. Większa grupa umożliwia nie tylko bardziej precyzyjne oszacowanie powtarzalności, ale także identyfikację tych czynników lub okoliczności, które pozostają w związku z niedostateczną powtarzalnością. Powtarzalność może dotyczyć odpowiedzi w skali ilościowej lub jakościowej (odpowiedzi w skali

półilościowej należy traktować w konwencji skali jakościowej).

Dla oceny zgodności w zakresie skali ilościowej często stosowano analizę korelacji liniowej. W tym przypadku wnioskowanie na podstawie wartości współczynnika korelacji liniowej 'r' jest błędne, ale warto przyrzeć się wykresowi korelacyjnemu. Może on ujawnić systematyczne zniekształcenie, obecne nawet przy dużych wartościach współczynnika 'r'. Właściwe metody oceny zgodności w tym przypadku obejmują opisową procedurę Bland-Altman'a i obliczenie tzw. korelacyjnego współczynnika zgodności Lin'a. Druga z metod umożliwia bezpośrednią słowną interpretację zgodności („słaba”, „średnia”, „znacząca”, „prawie całkowita”) na podstawie wartości współczynnika i jego 85% przedziału ufności.

W odniesieniu do skali jakościowej w ocenie zgodności stosuje się obliczenie odsetkowej zgodności (pozytywnej, negatywnej, ogólnej), test McNemar'a i - zapewne najczęściej - test Kappa (test Kappa Cohen'a). Popularność testu Kappa wynika z faktu, że poza określeniem statystycznej znamienności zgodności możliwe jest opisanie jej przy pomocy arbitralnej skali werbalnej (zwykle jest stosowana pięciostopniowa skala, od „bardzo słabej zgodności” do „bardzo dużej zgodności”). Metody oceny zgodności odpowiedzi jakościowych mają znaczenie uniwersalne - odpowiedzi w skali ilościowej można poddać transformacji do zmiennych jakościowych (np. na podstawie liczby przebytych ostrych zespołów wieńcowych można wyodrębnić badanych z „częstymi” i „rzadkimi” ostrymi zespołami wieńcowymi). Wymienione metody nie wyczerpują arsenału statystycznych metod oceny zgodności. Statystyki wyrażające poziom zgodności są liczne i można tu wspomnieć o współczynnikach Yule'a lub Cramer'a, czy współczynnika Kendall'a szczególnie przydatnym w przypadku skal półilościowych.

Trafność

Trafność kwestionariusza określa się na podstawie zgodności wyniku badania kwestionariuszowego z wynikiem badania referencyjnego, będącego w tym przypadku „złotym standardem”. Jest to uproszczone spojrzenie na omawiany atrybut kwestionariusza. W literaturze przedmiotu omawia się różne aspekty trafności kwestionariusza - między innymi trafność teoretyczną i treściową, trafność wewnętrzną i zewnętrzną. Skomplikowane podejście do zagadnienia trafności jest uzasadnione na gruncie rozważań teoretycznych, ale w praktyce trafność można opisać i zbadać w stosunkowo prosty i czytelny sposób. Założenie, że kwestionariusz jest narzędziem diagnostycznym sprawia, że kluczowe jest określenie w jakim stopniu jego stosowanie dostarcza wyników prawdziwie dodatnich,

prawdziwie ujemnych, fałszywie dodatnich i fałszywie ujemnych. Każda z wymienionych właściwości wymaga porównania wyniku badania kwestionariuszowego ze stanem faktycznym, określonym przez niezależne badanie, uznane za standardowe postępowanie. Przykład takiej procedury został wcześniej przedstawiony w odniesieniu do kwestionariuszowej diagnozy przewlekłego zapalenia oskrzeli. Zdarza się, że wartość diagnostyczną posiada pojedyncze pytanie. Na przykład odpowiedź na pytanie o intensywność palenia tytoniu można konfrontować z bezpośrednim pomiarem stężenia kotoniny w ślinie („złoty standard w tym przypadku”). Zwykle diagnoza konkretnego stanu wymaga uwzględnienia szeregu pytań i odpowiedzi. Dobrym przykładem jest próba rozpoznania astmy. W tym przypadku objawy choroby są określane na podstawie różnych pytań (np. przeszłe i aktualnie występujące napady duszności, świsty w klatce piersiowej, duszność wysiłkowa, kaszel nocny, objawy alergiczne itd.). Znaczenie poszczególnych odpowiedzi i/lub ich kombinacji określa się poprzez ich porównanie ze stanem faktycznym badanych, ustalonym w wyniku niezależnego, klinicznego postępowania diagnostycznego (z wynikiem astma = tak lub nie).

Dostępność wyniku badania referencyjnego („złotego standardu”) umożliwia zastosowanie konwencji tabeli czteropolowej dla obliczenia poszczególnych miar trafności (podstawowe są wymienione pod poniższą tabelą):

Wynik Badania Kwestionariuszowego	Wynik Badania Referencyjnego	
	+	-
↓ ↓ ↓ +	A	B
-	C	D

$$\text{Ogólna zgodność} = (a+d) / (a+b+c+d)$$

$$\text{Czułość} = a / (a+c)$$

$$\text{Swoistość} = d / (b+d)$$

$$\text{Diagnostyczny dodatni współczynnik
prawdopodobieństwa} = \text{czułość} / (1-\text{swoistość})$$

$$\text{Diagnostyczny ujemny współczynnik
prawdopodobieństwa} = (1-\text{czułość}) / \text{swoistość}$$

Procedura oceny trafności kwestionariusza jest analogiczna do procedury oceny trafności nowych testów diagnostycznych o znaczeniu klinicznym. W przypadku kwestionariusza duże znaczenie ma wybór takich pytań, których łączne znaczenie odznacza się maksymalną wartością poszczególnych wskaźników. Jest to pożądanym efektem standaryzacji - nawet gdy niemożliwy do osiągnięcia niezbędne jest poznanie trafności kwestionariusza, o ile to możliwe. Przedstawiona procedura dotyczy zewnętrznego wymiaru trafności („kwestionariusz a standard”). Przedtem jednakże warto zastanowić się nad jego wewnętrzną poprawnością. Pierwszym wymogiem, przedstawionym powyżej, jest dołożenie starań, aby

był on akceptowalny. Drugim wymogiem jest utrzymanie zdolności różnicującej poszczególnych pytań i ich zdolności do dostarczania uzupełniającej się wzajemnie informacji, istotnej w kontekście badanego zjawiska (np. w przypadku astmy tę właściwość posiadają pytania dotyczące, kolejno, występowania kaszlu, występowania świstów w klatce piersiowej, napadów duszności, napadów duszności ze świstami ...). Tę cechę kwestionariusza ocenia się zwykle obliczając tzw. współczynnik alfa Cronbach'a. Jego wartość odzwierciedla poziom korelacji pomiędzy odpowiedziami na poszczególne częściowe pytania, a łącznym wynikiem, ustalonym na podstawie odpowiedzi na wszystkie częściowe pytania. Taki zestaw pytań uważa się za przydatny, gdy wartość współczynnika Cronbach'a kształtuje się w zakresie 0,7 - 0,9 (w skali od 0,0 do 1,0). Małe wartości współczynnika sugerują, że poszczególne pytania dotyczą różnych problemów (np. w odniesieniu do astmy pytanie o kaszel może przede wszystkim „wychwytywać” przewlekłe zapalenie oskrzeli, a pytanie o napady duszności astmę). Maksymalne wartości współczynnika są z kolei sygnałem, że różne pytania mierzą praktycznie ten sam efekt (np. pytanie o napady duszności i pytanie o napady duszności ze świstami w klatce piersiowej) - w takiej sytuacji można wykluczyć jedno z pytań z zestawu stosowanego dla oceny trafności kwestionariusza i ponowić tę ocenę. Nie ma jednakże prostej wskazówki, jak dalece wykorzystywać procedurę Cronbach'a. Podane kryteria optymalnej wartości współczynnika alfa mają znaczenie arbitralne i nie powinny zdominować decyzji wynikających z wiedzy epidemiologicznej i medycznej oraz z celów badania, na potrzeby którego kwestionariusz został przygotowany.

Zalecenia praktyczne

Niniejszy esej oferuje autorskie spojrzenie na kwestię wiarygodności kwestionariuszy stosowanych w badaniach epidemiologicznych. To spojrzenie nie uwzględnia pełnej palety narzędzi walidacyjnych opisywanych w literaturze metodologicznej. Przedstawione procedury mają jednakże - zdaniem autora - duże praktyczne znaczenie. Podobnie, autorski charakter ma rozłożenie akcentów w ramach wielostopniowej procedury walidacyjnej, co wynika z kilku przesłanek:

Po pierwsze, wszyscy mamy prawo stosować autorskie kwestionariusze, ale pod warunkiem, że przeszły one procedurę walidacji.

Po drugie, korzystanie ze standardowych kwestionariuszy tłumaczonych z języka obcego powinno być także poprzedzone procedurą walidacyjną.

Po trzecie, nie wszystkie autorskie kwestionariusze muszą być poddane pełnej procedurze walidacyjnej. Ocena trafności jest potrzebna, gdy kwestionariusz ma służyć wykryciu konkretnego zjawiska zdrowotnego i możliwa, gdy to zjawisko można równolegle wykryć przy użyciu uznanej metody referencyjnej.

Zakres procedury walidacyjnej może być bardzo szeroki, ale zdaniem autora niezbędne jest uwzględnienie następujących elementów:

a) Ocena akceptowalności nie posiada standardowej procedury, ale powinna uwzględniać takie aspekty jak prostota, zrozumiałość i liczba pytań.

b) Ocena rzetelności jest oceną zgodności odpowiedzi. W odniesieniu do odpowiedzi ilościowych można wykorzystać procedurę Bland-Altman'a oraz korelacyjny współczynnik zgodności Lin'a. W odniesieniu do odpowiedzi jakościowych powinna obejmować obliczenie odsetkowej zgodności odpowiedzi oraz test Kappa.

c) Ocena trafności wymaga obliczenia, w stosunku do „złotego standardu”, ogólnej odsetkowej zgodności, czułości, swoistości, a także diagnostycznych współczynników prawdopodobieństwa.

Opisana procedura oceny wiarygodności jest prosta i wystarczająca. Szkopuł jednakże w tym, że jest ona rzadko stosowana. Takie przynajmniej wrażenie odnosi się z lektury prac epidemiologicznych, bazujących na badaniu kwestionariuszowym. Nie ma idealnych kwestionariuszy i naiwnością byłoby oczekiwać, że wyniki oceny ich wiarygodności sięgną maksymalnych wartości wskaźników. Nie zwalnia to od obowiązku poinformowania o wynikach walidacji i na to jest miejsce w naukowych opracowaniach danych pozyskanych w trakcie badań kwestionariuszowych. Postępowanie takie należy do kanonu dobrej praktyki epidemiologicznej.

(JEZ)

Redaktor
Jan E. Zejda
Przewodniczący Polskiego Stowarzyszenia Epidemiologicznego