

Piotr Samel-Kowalik, Aneta Tomaszewska, Bolesław Samoliński

REPORTING SURVEY PARTICIPATION RATES – TOWARD STANDARDIZATION IN EPIDEMIOLOGICAL STUDIES

WSKAŹNIKI REALIZACJI PRÓBY BADAWCZEJ – W KIERUNKU STANDARYZACJI RAPORTOWANIA W BADANIACH EPIDEMIOLOGICZNYCH

Medical University of Warsaw, Department of Prevention of Environmental Hazards and Allergology
Warszawski Uniwersytety Medyczny, Zakład Profilaktyki Zagrożeń Środowiskowych i Alergologii

ABSTRACT

Participation rates are recognized as key indicators of the quality of the data collected during a survey. This is particularly important in epidemiological studies where the findings are extrapolated from the study sample to the general population. The response rate is very frequently selected as the key indicator of the overall quality of the data collected but the literature search shows that there are many different ways to calculate it. Only the use of the whole set of participation rates allows assessment of both the quality of the survey (the quality of the data collected) and the reliability of the fieldwork. Standardization of the methods for calculating participation rates seems to be a prerequisite for a reliable comparison of the quality of population surveys.

Key words: *epidemiological study, participation rates, response rate, standardization*

STRESZCZENIE

Wskaźniki realizacji próby badawczej są jedną z kluczowych informacji, która wskazuje na jakość materiału zebranego w trakcie badania. Szczególne znaczenie ma to w badaniach epidemiologicznych, na podstawie których wyciąga się wnioski ekstrapolowane na populację generalną. W badaniach populacyjnych, podstawowym wskaźnikiem wybieranym przez badaczy do oszacowania stopnia realizacji próby jest wskaźnik *response rate*, a literatura tematu wskazuje na odmienność sposobów obliczania tego wskaźnika przez autorów prac. Jedynie wykorzystanie całego pakietu wskaźników pozwala na określenie nie tylko jakości badania (jakości zebranych danych), lecz także rzetelności prowadzonych prac terenowych. Standaryzacja metod obliczania wskaźników realizacji próby jest niezbędna dla możliwości rzetelnego porównywania jakości badań populacyjnych.

Słowa kluczowe: *badania epidemiologiczne, wskaźniki realizacji próby, wskaźnik odpowiedzi, standaryzacja*

INTRODUCTION

In population studies, estimation of the external¹ and internal² validity of a survey is based on the available information on how samples were drawn³

WPROWADZENIE

W badaniach populacyjnych oszacowanie wiarygodności zewnętrznej¹ i wewnętrznej² wymaga dostępności informacji o sposobie losowania³ oraz poziomie realizacji

¹ The degree to which the survey findings can be generalized beyond the survey sample, i.e. to the entire population from which the sample is derived or to other populations [1].

² The degree to which a measurement, questionnaire, test or a whole survey or any data set measures what it intends to measure [1].

³ Here, a method of drawing a random sample using probability sampling.

¹ Stopień, w jakim wyniki badania mogą mieć zastosowanie poza próbą użytą w badaniu, tzn. mogą być uogólnione na populację, z której pobrano próbę lub na inne populacje. (1)

² Stopień, w jakim pomiar, kwestionariusz, test lub całe badanie, lub jakikolwiek zbiór danych mierzy to, co w założeniu ma mierzyć. (1)

³ Tu rozumianego jako metodzie doboru jednostek losowania do badania, która oparta jest o metody probabilistyczne.

and the participation rate levels⁴. Assessment of the sampling method allows determining whether it correctly represents the target population while the participation rates indicate the validity⁵ of the findings. A survey sample is drawn from the source population (general population) and to ensure representative sampling the procedure should use the sampling frame. A survey sample consists of the representatives of the population to which the survey questions refer. In Poland, the PESEL or TERYT registers are used as the sampling frames for population studies. These registers allow selection of study groups based on predefined features.

Historical background

Although the *response rate* (RR) is the indicator most commonly selected in the evaluation of the quality of questionnaire surveys, we still lack a generally accepted comprehensive definition of the term and one way to calculate the rate. *Joy Williams-Jones* in the paper on the lack of agreement on the standardization of *response rate* terminology selected 29 different definitions specified out of 40 responses by personnel involved in the computation of *response rates* in institutions/companies which were commercial research suppliers. *Frederick Wiseman* and *Philip R. McDonald* sent to researchers raw data on refusal to participate, complete interviews, respondents who did not meet the eligibility criteria or respondents who were permanently inaccessible. The aim was to find out how the *response rate* was actually calculated. The achieved *response rate* was between 12% and 90% (2, 3).

Frederick Kviz in the 1970s was one of the first authors to attempt a systematic approach to calculation of *response rates* (4). In 1982, The Council of American Survey Research Organizations (CASRO) published guidelines on the recommended methods of calculating indicators of survey quality. Further studies in this area were published by *Judith Lessler* and *William Kalsbeck* (5), and *Robert M. Groves* (6). The report of The American Association for Public Opinion Research (AAPOR) (9th edition, 2016) offers guidelines on reporting survey-based findings.

Current standards

The guidelines and standards have been developed in the United States and because of differences in the research traditions, some methodological procedures and the ways of conducting surveys, some of these recommendations are not directly applicable to Europe. *Peter Lynn* and his coworkers from The

próby badawczej⁴. Ocena doboru próby pozwala odpowiedzieć na pytanie, czy odpowiednio reprezentuje ona populację docelową. Opis realizacji próby pozwala określić trafność⁵ pomiaru. Wybór badanej próby z populacji źródłowej (populacji generalnej) w celu zapewnienia reprezentatywności powinien być oparty o operat losowania. Próba składa się z przedstawicieli (reprezentantów) populacji odnośnie której chcemy pozyskać informacje. W praktyce badawczej w Polsce dla badań populacyjnych wykorzystuje się operat PESEL lub TERYT. Są to spisy jednostek, które umożliwiają dobranie próby badawczej za pomocą zdefiniowanych cech.

Tło historyczne

W badaniach kwestionariuszowych najczęściej wybieranym przez badaczy jest wskaźnik *response rate* (RR)⁶. Mimo wielu lat badań, w dalszym ciągu nie ma jednolitej i kompleksowej definicji tej miary. Badacze na wiele różnych sposobów próbują ten wskaźnik obliczać. *Joy Williams-Jones* w pracy poruszającej problem braku standaryzacji tego miernika wykazała, że wśród 40 respondentów (osób zajmujących się wyliczaniem RR w instytucjach i firmach prowadzących badania) funkcjonowało 29 definicji tego terminu. Z kolei *Frederick Wiseman* i *Philip R. McDonald* rozesłali do badaczy surowe dane zawierające informacje o odmowach udziału w badaniu, wywiadach zrealizowanych, respondentach niespełniających kryteriów włączenia do badania oraz trwałej niedostępności tych osób. Ich celem było zebranie danych dotyczących sposobu wyliczenia *response rate*. Z tego samego zestawu informacji otrzymali zwrótnie wyniki RR w przedziale od 12 do 90% (2,3).

Jednym z pierwszych badaczy, który podjął próbę usystematyzowania sposobów wyliczania *response rate* był *Frederick Kviz* w latach '70-tych XX wieku (4). W 1982 r. *Council of American Survey Research Organizations* (CASRO) przedstawiła rekomendacje dotyczące preferowanych sposobów szacowania wskaźników. W kolejnych latach tematyką zajmowali się *Judith Lessler* i *William Kalsbeck* (5) oraz *Robert M. Groves* (6). Jedną z nowszych prób aktualizacji wytycznych jest raport *American Association for Public Opinion Research* (AAPOR). Obecnie obowiązuje dziewiąta edycja wydana w 2016 r. (7).

Aktualne standardy

Powyższe wytyczne i standardy opracowane zostały w Stanach Zjednoczonych. Z uwagi na odmienną tradycję badawczych, niektórych aspektów metodologicznych i sposobów realizacji badań, nie wszystkie stosowane w nich zalecenia można wprost wykorzystać w warunkach

⁴ The percentage of originally selected subjects contacted or who completed the study.

⁵ Accuracy with which a selected tool measures what it intends to measure.

⁴ Odsetku przypadków wyznaczonych do badania, z którymi udało się nawiązać kontakt lub zrealizować badanie.

⁵ Trafność to dokładność, z jaką narzędzie mierzy to, co miało mierzyć.

⁶ Response rate (RR) – wskaźnik odpowiedzi.

Institute for Social and Economic Research (ISER) made recommendations for adapting the AAPOR guidelines to standard definitions of *response rate* in social surveys (8). These included the calculation of *response rates* in face-to-face interviews using name-based and address-based sampling, because the AAPOR guidelines refer mostly to address-sampling and techniques which traditionally have not been widely employed in Europe. Poland still lacks generally accepted consistent definitions of survey participation rates which could be used to standardize the calculation methods (9).

All authors, irrespective of the differences in the recommended definitions, advise the use, in addition to *response rates*, of some other measures such as *co-operation rate (COOP)*, *contact rate (CON)*, *refusal rate (REF)* or *eligibility rate (ELIG)* (2,7,8). The use of all these rates allows the evaluation of the quality of the data collected and of the reliability of the fieldwork. In order to compare the reliability of a number of surveys, each should be described using the same set of participation rates. When the information on the accepted definitions of terminology and the calculation method(s) is missing, it becomes difficult to evaluate the survey quality and interpret the findings. The *response rate* alone cannot be used to determine the extent of non-response error or to decide whether such error actually exists (7, 10, 11).

Groves advised to check the following survey design aspects to gain a better knowledge and understanding of the characteristics of a particular survey (6):

- Whether all sample units⁶ meet the eligibility criteria (it is possible that some units on the sampling frame are not members of the target population).
- Whether each sampling unit⁷ contains one sample element or many.
- Whether all sample units have the same probability of being selected.
- Whether substitution at the sampling stage is permitted.

Answers to these questions may aid in understanding the difference(s) in methodology between surveys with a potential impact on the calculation of survey participation rates. With a correct assessment of the above aspects it is easier to assess the eligibility of a sampling unit. Outcome categories are employed to collect information on conducting the survey and the participation rates. Knowledge of the eligibility established for each sampling unit allows detecting any sampling error and identifying its causes. The characteristic features of the sampling frame are another key factor. For example,

⁶ A sample unit is a sample element selected for participation in a survey, may be a person or a household.

⁷ A sampling unit is a sample element (a person or a household) selected for participation in a survey by randomization.

kach europejskich. Próby dostosowania wytycznych, dla których punktem wyjścia były rekomendacje AAPOR, dokonał Peter Lynn i jego współpracownicy z *Institute for Social & Economic Research (ISER)* (8). Dotyczyły one między innymi wykorzystania mierników w badaniach realizowanych techniką wywiadów bezpośrednich (*face-to-face*) na próbach imiennych i adresowych, ponieważ wytyczne AAPOR odnoszą się głównie do prób adresowych i technik, które nie mają tradycji w Europie. W Polsce w dalszym ciągu brakuje ogólnie przyjętych standardów, na podstawie których można dokonać ujednolicenia metod wyliczania wskaźników (9).

Wszyscy autorzy rekomendacji, niezależnie od różnic w propozycjach unifikacji definicji i pojęć, oprócz prezentacji samego RR zalecają również stosowanie kilku innych mierników, takich jak: *co-operation rate (COOP)*⁷, *contact rate (CON)*⁸, *refusal rate (REF)*⁹, *eligibility rate (ELIG)*¹⁰ (2,7,8). Wykorzystanie całego zestawu wskaźników pozwala na określenie nie tylko jakości zebranego materiału, lecz także rzetelności prowadzonych prac terenowych. Aby możliwe było porównanie wiarygodności badań każde powinno być opisane przy ich użyciu. Brak informacji o przyjętych definicjach i sposobie wyliczenia wskaźników prowadzi do trudności w ocenie jakości badania i w interpretacji wyników. Przedstawienie wyłącznie wartości *response rate* nie wystarcza do określenia, jak duży jest błąd wynikający z występowania braków odpowiedzi, ani czy w ogóle istnieje (7,10,11).

Aby lepiej poznać i zrozumieć charakterystykę konkretnego badania Groves sformułował następujące pytania (6):

- Czy wszystkie jednostki badania¹¹ spełniają kryteria włączenia? (niektóre jednostki, do których dołączymy, nie muszą pochodzić z populacji badanej)
- Czy każda jednostka losowania¹² zawiera jeden element próby, czy wiele?
- Czy wszystkie jednostki badania mają takie same prawdopodobieństwo wylosowania?
- Czy możliwe jest zastępowanie jednostek badania innymi na poziomie doboru do badania?

Mogą one pomóc w zrozumieniu odmienności metodologicznej badań, która może mieć znaczenie przy obliczaniu wskaźników realizacji próby. Właściwa ocena wskazanych w pytaniach zagadnień, pozwala między innymi na prawidłowe określenie statusu jednostki badania. Do gromadzenia informacji o przebiegu badania i stopniu wykorzystania próby stosuje się tak zwane *kategorie wynikowe*. Wiedza o tym, jaki status został w trakcie badania

⁷ *Co-operation rate* – wskaźnik współpracy.

⁸ *Contact rate* – wskaźnik kontaktu.

⁹ *Refusal rate* – wskaźnik odmowy.

¹⁰ *Eligibility rate* – wskaźnik określający przynależność do populacji badanej.

¹¹ Jednostka badania – wyznaczony do badania element próby, którym może być osoba lub gospodarstwo domowe.

¹² Jednostka losowania – element próby, którym może być osoba lub gospodarstwo domowe, wybrane do badania na drodze randomizacji.

the PESEL register often used in Poland in name-based sampling includes a legal address but not a current residential address. This may influence the classification of the sampling units, which cannot be contacted during the entire span of the survey as they do not actually live at the address indicated as legal.

Outcome categories

Lynn proposed standard outcome categories for social studies, which include temporary outcome categories (used during the fieldwork and reflecting the current status of a case) and final outcome categories (ultimately all cases must be assigned a final outcome code (8). Figure 1 shows the categorization of final outcome categories which masks a more complex description of outcomes employing temporary categories which may or even should to some extent differ depending on the type of survey and in each case ought to be precisely defined and stated explicitly. A comprehensive approach to describe the final disposition of cases and calculating outcome rates for surveys is presented in the *American Association for Public Opinion Research* (7). The work of Stadtmueller is an example of adapting the AAPOR recommendations for the context of a particular country (12) as is the paper by Grzeszkiewicz-Radulska who used the adaptation of the AAPOR final disposition codes to analyze the category of respondent non-contact in the surveys conducted by the Center for Public Opinion Research (CBOS) (9).

przypisany każdej jednostce losowania pozwala na określanie, czy próba zawiera błąd oraz na ocenę przyczyn takiego stanu. W tym kontekście kluczowe znaczenie ma również charakterystyka operatu losowania. Na przykład operat PESEL, często wykorzystywany w próbach imiennych w Polsce, zawiera informacje o adresie zameldowania, a nie zamieszkania. Fakt ten może mieć wpływ na to, jak klasyfikowane powinny być jednostki losowania, z którymi nie można nawiązać kontaktu w całym okresie trwania badania, ponieważ wskazane osoby nie zamieszkują pod podanym adresem.

Kategorie wynikowe

Lynn zaproponował kategoryzację wyników realizacji próby, która zawiera kategorie tymczasowe (wykorzystywane w trakcie prac terenowych) oraz kategorie finalne (ostatecznie wszystkie statusy klasyfikowane są do jednej z grup) (8). Na rycinie 1 przedstawiony został podział kategorii finalnych, który jest „maską” dla bardziej złożonego systemu zapisu z pomocą kategorii tymczasowych, które mogą, a nawet powinny być w pewnym zakresie odmienne dla różnych typów badań i każdorazowo jasno zdefiniowane. Szczegółowy opis dla konkretnych przypadków przedstawiony został w opracowaniu *American Association for Public Opinion Research* (7). Jako przykład adaptacji do warunków danego kraju może posłużyć opracowanie Stadtmueller (12). Grzeszkiewicz-Radulska na przykładzie badań *Centrum Badania Opinii Społecznej* przedstawiła adaptację schematów klasyfikacyjnych do głównych kategorii schematu AAPOR (9).

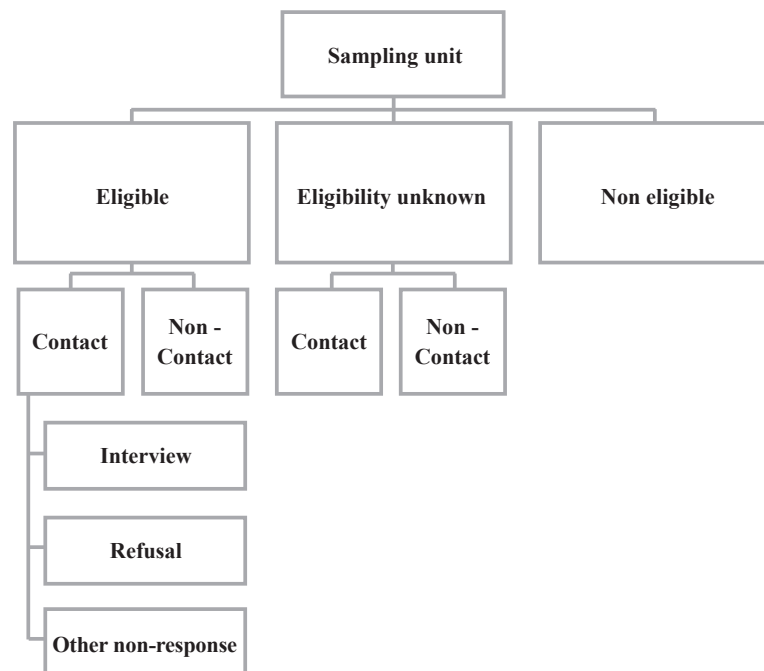


Fig. 1. Hierarchical arrangement of outcome categories: functional grouping of sampling units with events occurring during the fieldwork.

Ryc. 1. Hierarchiczny układ kategorii wynikowych – podział funkcjonalny grupujący jednostki losowania oraz związane z nimi zdarzenia występujące w trakcie realizacji próby badawczej.

Source: the diagram adapted from Lynn *et al.* (8, s. 7)

When the interviews are actually carried out, the events occurring during the course of the fieldwork may considerably differ between surveys depending on their design. Different temporary outcome categories apply to face-to-face interviews using name-based sampling, mail surveys using address-based sampling or online surveys.

Unified notation of the outcome codes may simplify reporting of survey findings (8):

- I – *complete interview*
- P – *partial interview*
- NC – *non-contact*
- R – *refusal*
- O – *other non-response*
- UC – *unknown eligibility, contacted*
- UN – *unknown eligibility, non-contact*
- NE – *not eligible*
- e_c – estimated proportion of contacted cases of unknown eligibility that are eligible
- e_N – estimated proportion of non-contacted cases of unknown eligibility that are eligible.

DEFINITIONS OF PARTICIPATION RATES

Considering that the *American Association for Public Opinion Research* guidelines are not fully applicable to surveys conducted in Europe, the following definitions of participation rates are based on the recommendations by *Lynn et al.* (8), who stated that they had taken the AAPOR standards as the starting point, but adapted them to the European survey context. The formulas shown in this paper are applicable to surveys using name-based and address-based sampling.

Participation rates serve two main purposes which have distinct implications for the definitions of these measures. First, the *response rate* is an indicator a survey (collected data) quality and should correctly reflect the structure of the survey population. In some instances, calculation of the *response rate* requires weighting by inverse selection probabilities, especially when stratified sampling is used. Second, the *response rate* may serve to assess the quality of the fieldwork and then unweighted *response rates* are more appropriate. The same practice of reporting applies to *cooperation rates* and *contact rates*.

Response rate (RR)

The *response rate* indicates how many interviews were achieved as the proportion of those selected (randomized) for the survey. In other words, it is the rate of complete or partial interviews to those selected for the survey.

$$RR = \frac{I + P}{(I + P) + (R + NC + O) + e_c UC + e_N UN} \quad [1]$$

W trakcie realizacji wywiadów prowadzonych z wykorzystaniem różnych technik, zdarzenia związane z realizacją wywiadów mogą znacznie się od siebie różnić. Inne kategorie tymczasowe stosuje się dla prób imiennych realizowanych techniką wywiadu bezpośredniego, a inne dla przeprowadzanych drogą pocztową na podstawie próby adresowej czy za pomocą Internetu.

W celu unifikacji wykorzystania kodów można stosować również ujednolicony sposób zapisu używanych skrótów (8):

- I – wywiad zrealizowany (*complete interview*),
- P – wywiad zrealizowany częściowo (*partial interview*),
- NC – brak kontaktu (*non-contact*),
- R – odmowa (*refusal*),
- O – inna przyczyna braku realizacji wywiadu (*other non-response*),
- UC – jednostka, o której nie wiadomo, czy należy do badanej populacji, z którą nawiązano kontakt (*unknown eligibility, contacted*),
- UN – jednostka, o której nie wiadomo czy należy do badanej populacji, z którą nie nawiązano kontaktu (*unknown eligibility, non-contact*),
- NE – jednostka która nie należy do badanej populacji (*not eligible*),
- e_c – szacowana proporcja jednostek należących do badanej populacji, z którymi nawiązano kontakt, ale dla których brak informacji o przynależności do badanej populacji,
- e_N – szacowana proporcja jednostek należących do badanej populacji, z którymi nie udało się nawiązać kontaktu.

DEFINICJE WSKAŹNIKÓW

Z uwagi na fakt, że wytycznych *American Association for Public Opinion Research* nie można we wszystkich przypadkach odnieść do warunków europejskich, przedstawione poniżej definicje wskaźników zostały oparte o opracowanie *Lynn'a* (8). Jak sam autor zaznacza terminologia stosowana w tej propozycji, wszędzie tam gdzie było to możliwe, została oparta o wytyczne AAPOR. Zaprezentowane wzory można zastosować do prób imiennych i adresowych.

Wskaźniki realizacji próby służą dwóm głównym celom, które mają odrębne implikacje dla definicji tych miar. Pierwszy to miara jakości badania – zebranych w nim danych – dlatego *response rate* powinien prawidłowo odzwierciedlać strukturę badanej populacji. Wymaga to, w niektórych przypadkach, zastosowania do jego obliczenia przeważenia grupy badanej według techniki *inverse selection probabilities*, która będzie użyteczna, jeśli dobór do badania miał charakter warstwowy. Drugi to ocena jakości pracy terenowej, dla której można używać nieprzeważonych miar. W takiej

The e_cUC and e_nUM parameters are an important feature of the calculation as they assume uncertain eligibility of cases (whether they belong to the survey population or not). They are included in the denominator when the sampling frame does not allow direct identification of these sampling units which meet the inclusion criteria. If the eligibility has been established, it is assumed that $e=1$. With name-based sampling and established eligibility, it is possible not to count UC and UN in the denominator.

It is also possible to calculate the *response rate* without partial interviews counted in the numerator. This is the *full response rate* while in the overall *response rate* both complete and partial interviews are counted.

Co-operation rate (COOP)

The *co-operation rate* indicates the number of achieved interviews with those selected for the survey as the proportion of those ever contacted during the data collection period.

$$COOP = \frac{I + P}{(I + P) + R + O + e_c UC} \quad [2]$$

Contact rate (CON)

The *contact rate* indicates the proportion of all cases reached by the interviewer irrespective of the presence or absence of the target respondent although verbal interaction to inform about the interview is required.

$$CON = \frac{(I + P) + R + O + e_c UC}{(I + P) + (R + NC + O) + e_c UC + e_n UN} \quad [3]$$

Refusal rate (REF)

This indicator is especially important considering the growing numbers of refusals to participate in surveys. The purpose of the *refusal rate* is to indicate the proportion of all estimated eligible cases that refuse.

$$REF = \frac{R}{(I + P) + (R + NC + O) + e_c UC + e_n UN} \quad [4]$$

The AAPOR guidelines present the methods of calculating the participation rates for telephone and online surveys although the authors warn that such calculations may be difficult or virtually impossible. Anyhow, in that context there are six variants of calculating *response rates*, three variants of calculating *refusal rates* and four variants of calculating co-operation rates.

EXAMPLE CALCULATIONS

Using the study Epidemiology of Allergic Diseases in Poland (ECAP) as an example we present the

postaci powinno się również prezentować *co-operation rate* oraz *contact rate*.

Wskaźnik odpowiedzi (response rate – RR)

Response rate wskazuje na to, jaka liczba wywiadów została zrealizowana w stosunku do tych wybranych (wylosowanych) do realizacji. Celem wskaźnika jest przedstawienie poziomu realizacji wyznaczonej do badania próby. Inaczej mówiąc, jest to stosunek liczby jednostek, z którymi zrealizowano wywiad w całości lub częściowo, do liczby jednostek wyznaczonych do badania.

$$RR = \frac{I + P}{(I + P) + (R + NC + O) + e_c UC + e_n UN} \quad [1]$$

Elementami równania, na które należy zwrócić uwagę, są parametry e_cUC oraz e_nUM , wnoszące do mianownika założenia wynikające z niepewności dotyczącej przynależności jednostki losowania do badanej populacji. Będą one miały zastosowanie, kiedy operat losowania nie będzie pozwalał na bezpośrednie wskazanie jednostek losowania spełniających kryteria włączenia. Jeśli jednak status każdego z przypadków został ustalony uprzednio, parametr e przyjmuje wartość 1. Dla prób imiennych, dla których uprzednio stwierdzona została przynależność do badanej populacji parametry UC i UN mogą być z mianownika usunięte.

Możliwe jest również obliczenie *response rate*, dla którego z licznika wykluczone zostaną wywiady niepełne. Wówczas taki wariant nazywany jest *pełnym response rate (full response rate)*, w odróżnieniu do *ogólnego response rate (overall response rate)*, który je zawiera.

Wskaźnik współpracy (co-operation rate – COOP)

Co-operation rate wskazuje liczbę zrealizowanych wywiadów z jednostkami wyznaczonymi do badania, w stosunku do tej liczby jednostek, z którymi skontaktowano się kiedykolwiek w trakcie gromadzenia danych.

$$COOP = \frac{I + P}{(I + P) + R + O + e_c UC} \quad [2]$$

Wskaźnik kontaktu (contact rate – CON)

Wynik *contact rate* wskazuje proporcję wszystkich jednostek, z którymi nawiązano jakikolwiek kontakt, niezależnie od obecności wyznaczonego do badania respondenta i informacji o możliwości przeprowadzenia z nim wywiadu.

$$CON = \frac{(I + P) + R + O + e_c UC}{(I + P) + (R + NC + O) + e_c UC + e_n UN} \quad [3]$$

methods of calculating particular participation rates (13). The survey sample was randomly selected and the PESEL register was the sampling frame. Each case (sampling unit) had an established and known eligibility. The survey was conducted as face-to-face interviews using computer-assisted personal interviewing (CAPI) technique.

Considering the nature of the study, in the calculation of the participation rates only cases belonging to the left arm (*eligible*) of the diagram in Fig. 1 were included. Interviews not achieved were described using one of the 11 causes listed in Table 1. A total of 20 455 interviews (I) were achieved from among 48 781 respondents ($I + R = NC + O$) randomly selected for participation in the survey. There were 11 284 refusals (R) and 17 042 cases which were not reached (NC). The UC, UN, e_c and e_N could be excluded from the calculations because all cases had established eligibility.

Table 1. Temporary disposition codes used to describe non-response and their conversion to final disposition codes.

Tabela 1. Kategorie tymczasowe służące do opisywania przyczyn braku realizacji wywiadów ankietowych oraz ich kategorie finalne

Reason for non-response (Temporary dispositions codes)	Final disposition codes	Number
The interviewer did not find anyone at the address	Non contact (NC)	8727
The respondent refused due to lack of time	Refusal (R)	1066
The respondent is temporarily unavailable at the address provided	Refusal (R)	2007
The respondent does not live at the designated address	Non contact (NC)	3203
The respondent is unavailable at the designated address (throughout the entire study period)	Non contact (NC)	3715
Refusal to let the interviewer into the apartment, housing estate, block of flats	Refusal (R)	1581
The respondent was indisposed, sick, intoxicated	Refusal (R)	296
The respondent refused because of the survey issues	Refusal (R)	1345
The respondent refused without giving a reason	Refusal (R)	4989
The interviewer did not find the address indicated	Non contact (NC)	241
The respondent reported as deceased or moved away	Non contact (NC)	1156

Response rate

According to the formula [1], to calculate the *response rate* the information on the number of accomplished interviews, refusals and instances of non-contact is required.

$$RR = \frac{20\,455}{20\,455 + 11\,284 + 17\,042} = 41.9\%$$

Co-operation rate

According to the formula [2], to calculate the *co-operation rate* the information on the number of accomplished interviews and of refusals is required.

Wskaźnik odmów (refusal rate – REF)

Miernik ten jest szczególnie istotny z uwagi na rosnące liczby odmów udziału w badaniach. Ma on na celu przedstawienie proporcji wszystkich wybranych do badania jednostek, które odmówiły udziału, do wszystkich pozostałych zdarzeń w badaniu.

$$REF = \frac{R}{(I + P) + (R + NC + O) + e_c UC + e_N UN} \quad [4]$$

W wytycznych AAPOR prezentowane są również sposoby obliczania wskaźników dla badań realizowanych z wykorzystaniem prób opartych o operaty telefoniczne oraz przez Internet, choć jak autorzy zaznaczają, wyliczanie wskaźników w tych przypadkach jest trudne, a czasami wręcz nawet niemożliwe. Możemy znaleźć różne warianty obliczania miar: sześć RR, trzy REF, cztery COOP.

PRZYKŁAD OBLICZANIA

Na przykładzie badania *Epidemiologia Chorób Alergicznych w Polsce* (ECAP) przedstawiony zostanie sposób wyliczenia poszczególnych wskaźników (13). Próba dobrana do badania miała charakter losowy, a jednostki losowane były na podstawie operatu PESEL i każda miała ustalony i znany status przynależności do populacji badanej. Badanie realizowane było metodą wywiadu bezpośredniego (*face-to-face*) z wykorzystaniem techniki CAPI (Computer assisted personal interviewing).

Z uwagi na charakter badania do wyliczania wskaźników przyjęte zostały wyłącznie przypadki znajdujące się

$$\text{CON} = \frac{20455 + 11284}{20455 + 11284 + 17042} = 65,1\%$$

Refusal rate

According to the formula [3], to calculate the *refusal rate* the information on the number of refusals, accomplished interviews and instances of non-contact is required.

$$\text{REF} = \frac{11\ 284}{20\ 455 + 11\ 284 + 17\ 042} = 23,1\%$$

SUMMARY AND CONCLUSIONS

Calculation and presentation of properly selected participation rates allows estimating the reliability of survey findings and a high *response rate* level is a good indicator. However, the *refusal rates* and *co-operation rates* help to estimate the degree of deficiency and its causes. This information also leads to a better understanding of the validity of the survey. This information examined in the context of particular research traditions and trends helps to decide whether the survey findings can be accepted as valid and extrapolated to larger populations. For instance, comparison of the participation rates calculated according to the same formulas for the ECAP study and for the CBOS surveys conducted in the same period of time revealed similar rates. Probably in both cases, the interviewers encountered similar problems in the course of fieldwork, i.e. increased migration of the population or increasing reluctance of people living in big cities to participate in such surveys.

The standardization of methods to calculate the participation rates is a prerequisite for a reliable comparison of the quality between population studies. It is especially important in the case of epidemiological studies where the findings are translated into policy decisions with impact on the organization and funding of the healthcare systems, including organized disease prevention measures. Still, it is difficult to specify the level of survey participation rate which could be recognized as the reference point. As shown by Mindell et al. (14) or Baruch and Holtom (15) participation in various population studies tends to decline and the mean *response rate* of less than 50% has become a rule in many countries. At the same time more and more studies are focused on elucidating the causes of non-response (16-21) since low *response rates* do not invariably suggest low quality and credibility of survey findings (11, 22).

REFERENCES

1. Słowniczek terminów epidemiologicznych - Epidemiological Review [Internet]. [cited 2019 Aug 30]. Available from: <http://www.przeglepidemiol.pzh.gov.pl/slowniczek-terminow-epidemiologicznych>

w lewym ramieniu (*eligible*) schematu przedstawionego na rycinie 1. Braki realizacji wywiadów były opisywane przy użyciu jednej z jedenastu przyczyn przedstawionych w tabeli 1. W badaniu zebrano 20455 (I) wywiadów spośród 48781 respondentów (I + R + NC + O) wylosowanych do badania. Odnotowano 11284 odmów udziału w badaniu (R) oraz 17042 przypadki, w których nawiązanie kontaktu nie było możliwe (NC). Ze względu na charakter badania parametry UC, UN, e_c oraz e_N można pominąć w obliczaniu wskaźników, ponieważ wszystkie jednostki losowania miały z góry ustalony status.

Response rate

Zgodnie ze wzorem [1] do wyliczenia tego wskaźnika potrzebne są informacje o liczbie zrealizowanych wywiadów oraz przyczynach braku realizacji wywiadów.

$$\text{RR} = \frac{20\ 455}{20\ 455 + 11\ 284 + 17\ 042} = 41,9\%$$

Co-operation rate

Zgodnie ze wzorem [2] do wyliczenia tego wskaźnika potrzebne są informacje o liczbie zrealizowanych wywiadów oraz liczbie odmów udziału w badaniu.

$$\text{COOP} = \frac{20\ 455}{20\ 455 + 11\ 284} = 64,4\%$$

Contact rate

Zgodnie ze wzorem [3] do wyliczenia tego wskaźnika potrzebne są informacje o liczbie zrealizowanych wywiadów, liczbie odmów udziału w badaniu oraz o liczbie braków realizacji wywiadów.

$$\text{CON} = \frac{20455 + 11284}{20455 + 11284 + 17042} = 65,1\%$$

Refusal rate

Zgodnie ze wzorem [4] do wyliczenia tego wskaźnika potrzebne są informacje o liczbie odmów udziału w badaniu, liczbie zrealizowanych wywiadów oraz o liczbie braków realizacji wywiadów.

$$\text{REF} = \frac{11\ 284}{20\ 455 + 11\ 284 + 17\ 042} = 23,1\%$$

PODSUMOWANIE I KONKLUZJE

Wyliczenie i prezentacja właściwie dobranych wskaźników ułatwia oszacowanie wiarygodności badania. Wysoka wartości wskaźnika *response rate* jest dobrym prognozykiem. Jednak informacja dotyczących odsetka odmów czy odsetka osób, z którymi nawiązano kontakt spośród tych wylosowanych do badania pozwala odpowiedzieć na pytanie jak duże odnotowano braki i określić z czego one

2. Shaw M, Bednall D, Hall J. A Proposal for a Comprehensive Response-Rate Measure (CRRM) for Survey Research. *J Mark Manag.* 2002 Jun;18(5–6):533–554.
3. De Leeuw ED. To mix or not to mix data collection modes in surveys. *J Off Stat.* 2005;21(5).
4. Kviz FJ. Toward a Standard Definition of Response Rate. *Public Opin Q.* 1977;41(2):265–7.
5. Lessler JT, Kalsbeek WD. Nonsampling error in surveys. New York: Wiley; 1992. 412 p. (Wiley series in probability and mathematical statistics).
6. Groves RM. Survey errors and survey costs. Hoboken, N.J: Wiley; 2004. 136 p. (Wiley series in survey methodology).
7. Standard Definitions - AAPOR [Internet]. [cited 2019 Apr 10]. Available from: [https://www.aapor.org/Standards-Ethics/Standard-Definitions-\(1\).aspx](https://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx)
8. Lynn P, Beerten R, Laiho J, Martin J. Recommended Standard Final Outcome Categories and Standard Definitions of Response Rate for Social Surveys [Internet]. Institute for Social and Economic Research; 2001 [cited 2019 Apr 10] p. 39. Available from: <https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2001-23.pdf>
9. Grzeszkiewicz-Radulska K. Jak w Polsce rozwija się zjawisko niedostępności respondentów? Analiza na przykładzie badań CBOS 1993–2011. *Przegląd Socjol.* 2012;61(1):83–111.
10. Sandler DP. On revealing what we'd rather hide: the problem of describing study participation. *Epidemiology.* 2002;13(2):117.
11. Morton LM, Cahill J, Hartge P. Reporting Participation in Epidemiologic Studies: A Survey of Practice. *Am J Epidemiol.* 2006 Feb 1;163(3):197–203.
12. Stadtmueller S. Adaptation of the AAPOR Final Disposition Codes for the German Survey Context. *Gesis Surv Guidel* [Internet]. 2019 [cited 2019 Aug 27]; Available from: <https://www.gesis.org/gesis-survey-guidelines/operations/adaptation-of-the-aapor-final-disposition-codes-for-the-german-survey-context/>
13. Samoliński B, Raciborski F, Lipiec A, Tomaszewska A, Krzych-Falta E, Samel-Kowalik P, et al. Epidemiologia Chorób Alergicznych w Polsce (ECAP). *Alergol Pol - Pol J Allergol.* 2014 Jan 1;1(1):10–8.
14. Mindell JS, Giampaoli S, Goesswald A, Kamtsiuris P, Mann C, Männistö S, et al. Sample selection, recruitment and participation rates in health examination surveys in Europe – experience from seven national surveys. *BMC Med Res Methodol.* 2015 Oct 5;15(1):78.
15. Baruch Y, Holtom BC. Survey response rate levels and trends in organizational research. *Hum Relat.* 2008 Aug;61(8):1139–60.
16. Abrahamsen R, Svendsen MV, Henneberger PK, Gundersen GF, Torén K, Kongerud J, et al. Non-response in a cross-sectional study of respiratory health in Norway. *BMJ Open.* 2016 Jan 6;6(1):e009912.
17. Cohen G, Duffy J. Are nonrespondents to health surveys less healthy than respondents? *J Off Stat.* 2002 Jan 1;18:13–23.
18. Drivsholm T, Eplöv LF, Davidsen M, Jørgensen T, Ibsen H, Hollnagel H, et al. Representativeness in population-based studies: a detailed description of non-response in a Danish cohort study. *Scand J Public Health.* 2006;34(6):623–31.
19. Knapstad M, Löve J, Holmgren K, Hensing G, Overland S. Registry-based analysis of participatory representativeness: a source of concern for sickness absence research? *BMJ Open.* 2016 21;6(10):e012372.
- wynikają. Dzięki temu możemy lepiej zrozumieć jaka jest wiarygodność. Osadzenie tych informacji w kontekście tradycji i trendów badawczych pozwala ocenić czy prezentowane wyniki można uznać za trafne oraz czy można je ekstrapolować na populację. Przykładowo porównanie wskaźników realizacji próby badawczej (wyliczonych według takich samych wzorów) badania ECAP ze wskaźnikami z badań CBOS, które realizowane były w podobnym okresie pokazuje, że osiągnięto zbliżone wartości. Z uwagi na okres w jakim prowadzono te badania, w trakcie prac terenowych prawdopodobnie natknięto się na problemy związane ze wzrostem migracji ludności lub wzrostem niechęci do udziału w tego typu badaniach wśród mieszkańców dużych miast.
- Standaryzacja metod obliczania wskaźników realizacji próby jest niezbędna dla możliwości rzetelnego porównywania jakości badań populacyjnych. Szczególnie istotne jest to w badaniach epidemiologicznych, których wyniki przekładają się na decyzje dotyczące organizacji i finansowania ochrony zdrowia oraz działań profilaktycznych. Jednocześnie ciężko jest mówić o normach, które należy przyjąć i uznać za referencyjne dla wskaźników realizacji próby. Jak pokazują prace Mindell (14) czy Baruch i Holtom (15) odnotowuje się coraz niższe wskaźniki realizacji w różnych typach badań populacyjnych, a wartości średnie *response rate* poniżej 50% w wielu krajach stały się regułą. Jednocześnie coraz większą wagę przykładają się do zrozumienia tego, jakie są przyczyny braku realizacji wywiadów (16–21), ponieważ niskie wartości wskaźnika *response rate* nie zawsze oznaczają niską jakość i wiarygodność (11, 22).

20. Knudsen AK, Hotopf M, Skogen JC, Overland S, Mykletun A. The health status of nonparticipants in a population-based health study: the Hordaland Health Study. *Am J Epidemiol.* 2010 Dec 1;172(11):1306–14.
21. Helakorpi S, Mäkelä P, Holstila A, Uutela A, Vartiainen E. Can the accuracy of health behaviour surveys be improved by non-response follow-ups? *Eur J Public Health.* 2015 Jun 1;25(3):487–90.
22. Galea S, Tracy M. Participation Rates in Epidemiologic Studies. *Ann Epidemiol.* 2007 Sep 1;17(9):643–53.

Received: 3.06.2019

Accepted for publication 16.09.2019

Otrzymano: 3.06.2019 r.

Zaakceptowano do publikacji: 16.09.2019 r.

Address for correspondence:

Adres do korespondencji:

Piotr Samel-Kowalik

Zakład Profilaktyki Zagrożeń Środowiskowych i Alergologii Warszawski Uniwersytet Medyczny
e-mail: piotr.samel@wum.edu.pl