

Andrzej Zieliński

ANALITYCZNE BADANIA KOHORTOWE

Badania kohort stanowią podstawowy rodzaj analitycznych badań epidemiologicznych. Celem tych badań jest zmierzenie i zazwyczaj porównanie zachorowalności w dwu lub wielu kohortach. W epidemiologii słowo kohorta oznacza grupę ludzi dzielących to samo doświadczenie lub cechę. Badanie może dotyczyć próby pobranej z całej populacji zamieszkującej na danym terenie lub jej odpowiednio zdefiniowanej subpopulacji. Taka subpopulacja może być określona przez cechy uczestników badania, takie jak wiek, płeć, grupa etniczna, zawód, wykształcenie, czy inne odpowiednio zdefiniowane wyróżniki. Na przykład kohorta wegetarian ścisłych obejmuje ludzi, którzy nie jedzą produktów zwierzęcych, kohortę urodzeń spaja ten sam rok lub okres urodzenia, a kohortę palaczy apierosów definiuje występujący u jej członków zwyczaj palenia papierosów.

Istotny z punktu widzenia rodzaju badania jest podział kohort na otwarte (*open*) oraz ustalone (*fixed*) i zamknięte (*closed*). Kohortę otwartą stanowi zbiór osób zdefiniowany na podstawie określonego stanu, w czasie trwania tego stanu. Uczestnictwo w kohorcie otwartej nie jest trwałe, lecz związane z definiującym je stanem. Na przykład w badaniu wpływu leków niesterydowych przeciwzapalnych na występowanie krwawienia z przewodu pokarmowego dana osoba należy do kohorty narażonych w czasie, gdy przyjmuje te leki, a do kohorty nienarażonych, gdy ich nie przyjmuje. Do otwartych należą też kohorty definiowane przez zamieszkanie na danym terenie.

Uczestnictwo w ustalonych i zamkniętych kohortach jest definiowane przez zdarzenie lub stan trwale związany z daną osobą, np. płeć, data urodzenia, odbycie służby wojskowej, przebycie określonej choroby. Z takiej kohorty nie można przejść do innej, ale w przypadku kohorty ustalonej można wypaść z obserwacji. Przykładem kohort

ustalonych są kohorty tworzone w eksperymentalnych badaniach klinicznych. W kohortcie zamkniętej, w założeniu wszyscy jej uczestnicy pozostają w obserwacji przez cały czas trwania kohorty. Miarą stosowaną w kohortach otwartych i ustalonych jest gęstość zapadalności, czyli stosunek liczby nowych zachorowań do sumy osobo-lat obserwacji, a w kohortach zamkniętych średnia zapadalność, czyli stosunek liczby nowych zachorowań do liczebności kohorty z uwzględnieniem czasu jej obserwacji.

Badanie kohortowe może dotyczyć jednej tylko kohorty i mieć charakter czysto opisowy. Wynikiem obserwacji prowadzonej przez określony przeciąg czasu jest wtedy zapadalność na określone choroby lub zmiany w przebiegu określonych stanów. Przykład badania opisowego opartego na jednej kohorcie stanowi rutynowy nadzór epidemiologiczny.

W kohortowym badaniu analitycznym występują dwie kohorty lub więcej: kohorta referencyjna składająca się z osób nienarażonych (nie poddanych ekspozycji) oraz jedna lub więcej kohort osób narażonych czyli posiadających cechę będącą przedmiotem badania pod kątem możliwego związku tej cechy z wystąpieniem choroby lub innego wyniku związanego ze zdrowiem. W rozdziale tym przedstawiony zostanie wstępny zarys układu badawczego kohort oraz elementy analizy wyników z uwzględnieniem niektórych źródeł błędów i sposobów ich korekty.

Układ badań kohortowych stosowany jest zarówno w kohortach obserwacyjnych, w których narażenie występuje w sposób niezależny od prowadzącego badanie, jak i w kontrolowanych badaniach klinicznych, należących do badań eksperymentalnych. Jednak ze względu na metodologiczną specyfikę badań klinicznych są one klasyfikowane i opisywane w podręcznikach osobno, a określenie „badanie kohortowe” używane jest niemal wyłącznie w stosunku do kohort obserwacyjnych.

Badanie kohortowe służy ocenie zdarzeń występujących w określonej populacji w określonym czasie. Rzadko przedmiotem badania jest cała populacja, choć obowiązkowy, powszechny system zgłoszeń nowo rozpoznanych chorób w ramach nadzoru epidemiologicznego można uznać za opisowe badanie kohortowe całej populacji poddanej nadzorowi. Jednak w ogromnej większości badań kohortowych kohorta stanowi

próbę pobraną z populacji na podstawie odpowiednio zdefiniowanej charakterystyki jej uczestników.

Dostosowanie miar zapadalności i oceny narażenia do typu kohorty

Ryzyko (zachorowalność zbiorcza) jest mierzone jako proporcja w odniesieniu do mianownika, którym jest liczebność kohorty, podczas gdy gęstość zachorowalności ma jako mianownik osobo-czas, na przykład osobo-lata, co daje dużo większą elastyczność w analizie wyników. Podczas gdy badania nastawione na obliczanie zachorowalności zbiorczej są konceptualnie związane z wydzieleniem kohort o specyficznej charakterystyce, badania w których mierzona jest gęstość zachorowalności nie zawsze są przywiązane do pojedynczych kohort. Dana osoba może mieć liczony swój osobo-czas w kilku grupach ekspozycji, ponieważ jej czasy obserwacji różnych ekspozycji mają swe indywidualne zaklasyfikowania. Na przykład osoba włączona do badań analizujących czynniki ryzyka choroby wieńcowej może udzielać swego osobo-czasu próbom, w których ryzyko jest definiowane jako nadciśnienie tętnicze, poziom trójglicerydów, ciężar ciała oraz poziom cholesterolu w osoczu. W każdej z tych grup czas obserwacji może być inny. Terminy otwarta lub dynamiczna kohorta określają zmienny zbiór ludzi włączanych do badań z różnych list sporządzonych według ekspozycji, gdzie jedna osoba może być na wielu listach. Na przykład mieszkańcy jakiegoś miasta stanowią kohortę otwartą, ponieważ zbiór jego mieszkańców jest zmienny, a różne narażenia mogą występować u tej samej osoby partycypującej w badaniach określonej choroby.

Jeżeli grupy ekspozycji w badaniu kohort są ustalone na początku obserwacji, i nie następuje przemieszczanie się osób badanych z jednej grupy do drugiej, mamy do czynienia z kohortą stacjonarną. Jeżeli w kohorcie stacjonarnej w czasie obserwacji niektóre osoby wypadają z kohorty lub występują w niej konkurencyjne narażenia, gęstość zachorowalności wciąż może być mierzona bezpośrednio i używana do oszacowania zachorowalności kumulatywnej i średniego czasu zachorowania z zastrzeżeniem, że ubytki z kohorty stanowią źródło błędu. Jeżeli w kohorcie stacjonarnej nie występują ubytki osób obserwowanych, kohorta stacjonarna spełnia kryteria kohorty zamkniętej. W takiej kohorcie możliwy jest pomiar zachorowalności bezwarunkowej

(*unconditional risk*) oraz średniego czasu przeżycia. Zagrożenia konkurencyjne, inne od badanego, mogą być powodem zawyżenia ich wyników.

W najprostszym układzie kohorty przyjmowane jest założenie, że narażenie jest ciągłe, niezmiennie co do intensywności i możliwe do identyfikacji. Umożliwia to zadanie podziału osób włączonych do badań na narażonych i nienarażonych. Taki ideał osiągnąć jest rzadko.

Określenie ekspozycji przewlekłej jest dużo trudniejsze niż ekspozycji w jakimś punkcie czasowym. Musimy uzmysłwić sobie okres czasu, w którym ekspozycja kumuluje się w dostatecznym stopniu, aby wywołać proces przyczynowo-skutkowy. Ta kumulacja może być złożoną funkcją intensywności ekspozycji i czasu. Czas indukcji teoretycznie rozpoczyna się w momencie uruchomienia procesu przyczynowo-skutkowego, ale projektując badania nie wiemy często, czy w ogóle jakiś związek przyczynowy między ekspozycją i chorobą ma miejsce, a zupełnie wyjątkowo możemy orzec, kiedy ekspozycja stanowi przyczynę wystarczającą. Epidemiolodzy przemysłowi i środowiskowi często liczą czas od pierwszej chwili ekspozycji, nie zakładając nawet, że pierwszy kontakt jest wystarczający do wywołania efektu, ale zakładając rolę początku ekspozycji w jej skumulowanym efekcie. W każdym projekcie badań sposób pomiaru czasu ekspozycji powinien być precyzyjnie podany. Na etapie planowania badań analiza ekspozycji powinna obejmować określenie liczby kategorii ekspozycji i ustalenie możliwie dokładnej definicji każdej z kategorii. Ekspozycja może być mierzona samym czasem jej trwania, czasem trwania pomnożonym przez średnią intensywność, maksymalną dawką (stężeniem) czynnika narażenia lub całkowitą sumą skumulowanej dawki. Na przykład ekspozycja palacza papierosów może być mierzona ilością lat nałogu, ilością „paczek-lat” (przez ile lat badany wypalałby jedną paczkę dziennie), lub oszacowaną ilością wypalonych papierosów. „Paczka-rok” to w perspektywie ponad czteroletniej 7305 papierosów wypalanych rocznie. Miary czasu-intensywności zakładają addytywność ekspozycji, a mianowicie to, że przyrost ekspozycji o określoną ilość jednostek da podobny efekt, niezależnie od tego, jak układa się w tych jednostkach proporcja czasu i intensywności. Na przykład pół paczki papierosów wypalanych codziennie przez 40 lat da tyle samo „paczek-lat” co dwie paczki dziennie wypalane

przez dziesięć lat. W rzeczywistości trzeba by dopiero wykazać, że obie te ekspozycje stanowią takie samo ryzyko choroby. Poza tym, nawet gdyby stanowiły to samo ryzyko raka płuc, nie muszą stanowić tego samego ryzyka choroby wieńcowej. Tego typu badania zostały wykonane i okazało się, że przy tej samej liczbie wypalonych papierosów waga czasu ekspozycji jest większa niż waga dziennej dawki.

Osobne i metodologicznie trudne zagadnienie odnośnie kohort dynamicznych stanowi liczenie czasu bez ekspozycji u osób narażonych. Jeśli latencja jest krótka, jak w przypadku przyjmowania niektórych leków, jako czas ekspozycji liczymy tylko czas przyjmowania tych leków, a pozostały czas u tych samych osób jest okresem bez ekspozycji. W przypadkach dłuższych czasów latencji, jej okres powinien być liczony jako czas bez ekspozycji, szczególnie tam, gdzie narażenie poprzedza efekt o czas krótszy od czasu latencji. Efekt, który wystąpił po początku narażenia w czasie krótszym od okresu latencji powinien być traktowany jako efekt innych, niebadanych przyczyn, a osoby których to dotyczy, nie powinny być uwzględniane w analizie jako przypadki.

W niektórych badaniach kohort ważne bywa ustalenie zależności pomiędzy dawką i odpowiedzią. W tych przypadkach konieczne jest posługiwanie się skalą ciągłą lub kategoryzacją ekspozycji, czyli przypisanie jej wartości liczbowych dających się ułożyć w wyczerpujący i rozłączny ciąg uporządkowany. Znaczy to że: 1) każda wartość ekspozycji jest przyporządkowana jakiejś kategorii, 2) nie ma takiej wartości zmiennej, która by była przypisana więcej niż jednej kategorii, 3) każda kolejna wartość charakteryzująca sąsiednią kategorię jest stale większa (lub mniejsza) od poprzedniej. Niekiedy jednak interesują nas różnice efektów pomiędzy kategoriami ekspozycji bez porównywania wielkości tych kategorii i wtedy dzielimy je na wyczerpujące i rozłączne kategorie nominalne, jak na przykład płeć. Zazwyczaj takie kategorie są trwale przypisane poszczególnym osobnikom i nie stwarzają problemu przy projektowaniu badań w kohortach statycznych. Natomiast kategorie ekspozycji definiowane przez zmienne kardynalne stwarzają problem w analizie wyników z kohort statycznych, gdyż przejście z jednej kategorii do drugiej wiąże się w tym wypadku ze zmianą kohorty, co stanowi poważne źródło błędów w analizie danych. W kohorcie dynamicznej nie stanowi to

tak wielkiego problemu, a jeżeli latencja nie jest bardzo długa. Dana osoba może mieć kolejno mierzony osobo-czas dla różnych kategorii, do których aktualnie należy.

Kohorty specyficznej ekspozycji i kohorty populacji ogólnej

W przypadku rzadkich ekspozycji, a w szczególności związanych z wykonywanym zajęciem, losowy dobór kohort wymagałby przebadania wielkiej liczby osób i dlatego optymalnym rozwiązaniem jest dobór kohorty narażonych wśród wybranych grup o znanym narażeniu. Przykładem takiej kohorty jest badanie skutków zdrowotnych „Agent Orange” w Wietnamie. O atrakcyjności badania tego rodzaju kohort stanowi też możliwość sprawdzania wielu efektów zdrowotnych wynikających z jednego rodzaju ekspozycji. Obserwacja z punktu widzenia pojedynczej ekspozycji została w tych badaniach rozszerzona na więcej niż jeden rezultat i badano w nich osobno wczesne i późne następstwa ekspozycji.

Kohorty populacji ogólnej są definiowane najczęściej terytorialnie jako mieszkańcy określonego obszaru administracyjnego np. gminy, miasteczka, czy dzielnicy. Następnie definiowane bywają ekspozycje i wszyscy uczestnicy kohorty są obserwowani pod kątem rozwijania się chorób będących przedmiotem zainteresowania. Oczywiście taka kohorta może być analizowana tylko jako kohorta dynamiczna, choćby ze względu na mobilność mieszkańców badanej miejscowości. Najbardziej znanym przykładem takiej kohorty jest Framingham Heart Study. W tym niewielkim miasteczku położonym kilkanaście mil na zachód od Bostonu epidemiolodzy i klinicyści z Boston University oraz z Narodowych Instytutów Zdrowia, pomiędzy rokiem 1950 a 1952 zidentyfikowali i przebadali 5127 mężczyzn i kobiet w wieku 30-59 lat, którzy byli wolni od chorób serca. Informacje zostały zebrane na temat zmiennych demograficznych, historii chorób przebytych, palenia papierosów oraz szeregu parametrów klinicznych i laboratoryjnych, jak tętniczne ciśnienie krwi, poziomy glikemii i inne. Co ciekawe, poziomy lipidów zaczęto oznaczać później, kiedy uzyskano dane z innych badań o ich znaczeniu dla rozwoju choroby wieńcowej. Członkowie kohorty byli, a niektórzy z nich są do tej pory obserwowani i badani w regularnych odstępach czasowych i monitorowani pod kątem rozwoju chorób serca. Na podstawie tych badań stało się możliwe zidentyfikowanie licznych czynników ryzyka choroby wieńcowej, a wśród nich: palenia papierosów, poziomu lipidów z rozpisaniem

ich na frakcje oraz nadciśnienia tetniczego. Liczba istotnych publikacji pochodzących z badań Framingham przekracza znacznie tysiąc, a do kohorty włączani są mieszkańcy miejsczka należący już do następnych dwóch pokoleń.

Kohorty prospektywne i retrospektywne

Badania kohort mogą być zarówno prospektywne jak i retrospektywne w zależności od relacji czasowej pomiędzy początkiem badania a wystąpieniem choroby. W obu typach kohort dobór uczestników badań musi być oparty na obecności lub nieobecności ekspozycji. W retrospektywnej kohorcie jednak zarówno ekspozycja jak i choroba musiały pojawić się przed rozpoczęciem badania. W prospektywnych kohortach ekspozycja może nastąpić przed rozpoczęciem badania, ale choroba, czy inny wynik tej ekspozycji (*outcome*) musi nastąpić dopiero w trakcie obserwacji, a więc już po rozpoczęciu badań. W niektórych wypadkach badanie kohort bywa dwukierunkowe; dane w tej samej kohorcie są zbierane zarówno prospektywnie, jak i retrospektywnie. Taki układ badań jest najbardziej użyteczny dla ekspozycji, które mają zarówno krótkotrwałe, jak i długotrwałe efekty. Jako przykład można wymienić substancję, która zwiększa ryzyko uszkodzeń płodów w ciągu roku lub dwu, a ryzyko raka w perspektywie dwóch dekad. Badania efektów dioksyn (*Agent Orange*) zostały przeprowadzone na personelu amerykańskiego lotnictwa wojskowego. Populację badaną reprezentowała próba 1264 osób, które uczestniczyły w stosowaniu tego defoliantu w Wietnamie pomiędzy 1962 a 1971 rokiem oraz próba 1264 osób personelu latającego, które odbywały misje transportu towarów w tym samym okresie, a z dioksynami nie miały kontaktu. Dane te były analizowane retrospektywnie, aby porównać problemy zdrowotne eksponowanych i nieeksponowanych w kategoriach tych efektów, które mogły być oszacowane we względnie krótkim czasie po ekspozycji, jak zmiany dermatologiczne, bezpłodność, wady wrodzone dzieci uczestników kohorty, uszkodzenia wątroby i zaburzenia psychiczne. Ze względu na poważne możliwości wzrostu ryzyka chorób nowotworowych kohorta była następnie obserwowana w prospektywnej części badań, aby umożliwić wykrycie efektów o długim czasie latencji.

Wybór układu kohorty prospektywnej lub retrospektywnej dla sprawdzenia naszej hipotezy winien opierać się na szeregu przesłanek merytorycznych i logistycznych. W

pierwszym rzędzie wybór ten zależy od dostępnego już materiału oraz możliwości doboru próby i przeprowadzenia własnych badań. Zwykle badania retrospektywne mogą być przeprowadzane szybciej i taniej niż badania prospektywne, ponieważ w momencie rozpoczynania badań wszystkie istotne zdarzenia już nastąpiły. Zatem układ retrospektywny kohort jest szczególnie efektywny w badaniach zjawisk o długich okresach latencji, wymagających wielu lat do osiągnięcia punktów końcowych. Jednak ze względu na to, iż dokumentacja do takich badań bywa prowadzona na wiele lat przed podjęciem badań, zależą one od jakości dokumentacji medycznej, rzadko reprezentującej odpowiedni poziom dokładności z punktu widzenia danego badania, choćby dlatego, że gromadzona ona była w innych celach niż to badanie. W badaniach prospektywnych odwołujemy się do znacznie świeższej dokumentacji, a w wielu przypadkach mamy okazję planowania i nadzorowania treści i sposobów gromadzenia danych. Zatem zysk w postaci mniejszych kosztów badań retrospektywnych winien być rozważany w świetle jakości dostępnej w tych badaniach informacji. Dobrze przeprowadzone prospektywne badanie kohort jest uznawane za szczególnie wartościowe źródło informacji epidemiologicznej.

Projektowanie badań kohortowych

W swej istocie badanie kohortowe stanowi podstawową procedurę sprawdzenia hipotezy, czy istnieje związek między określoną zmienną analizy, stanowiącą charakterystykę ekspozycji a zmienną będącą wskaźnikiem choroby lub innego stanu związanego ze zdrowiem, a będącego przedmiotem badania. Optymalny wynik dociekań badacza polegałby na znalezieniu związków przyczynowych między narażeniem i chorobą, czyli określenia tych narażeń, które stanowią podstawę ustalenia warunków koniecznych i wystarczających do wystąpienia choroby. Badania kohortowe mogą dać podstawy do tego rodzaju wnioskowań, ale ich bezpośredni wynik stanowi jedynie określenie związków statystycznych, jakie można ustalić w badanej próbie między tymi zmiennymi.

Jakość doboru próby, jej odpowiednia liczebność oraz jakość analizy statystycznej wpływają na znamienność wyniku, jak i na moc badania kohortowego jako testu sprawdzającego założone hipotezy. Im poprawniej jest przeprowadzane badanie: w jakim

zakresie jest ono wolne od błędów doboru próby, błędów zaklasyfikowania/pomiaru narażenia i choroby oraz od czynników zakłócających, tym mocniejszą może stanowić podstawę do dalszych rozumowań o charakterze przyczynowym. Dlatego w projektowaniu badania należy rozpocząć działanie od jasnego określenia celu badania oraz populacji, której ma ono dotyczyć i sformułowania hipotez, które badanie ma sprawdzić. Do tych założeń należy następnie dostosować jego układ i stosowane procedury w zakresie gromadzenia danych i ich opracowania statystycznego.

Pierwszym warunkiem jest właściwe zdefiniowanie badanych zmiennych dotyczących zarówno narażenia, jak i choroby oraz określenie populacji osób podatnych (*population at risk, vulnerable population*). Precyzja w definiowaniu narażenia i choroby zabezpiecza porównywalność ze sobą uczestników kohorty. Zaś warunek, że w kohorcie występują osoby podatne na badaną chorobę, czyli takie, u których prawdopodobieństwo wystąpienia tej choroby nie jest zerowe, jest oczywisty. W badaniu czynników ryzyka wystąpienia raka trzonu macicy nie mogą występować mężczyźni ani kobiety po histerektomii. Ale też moc testów służących do sprawdzania hipotez zależy od odpowiednio dobranego przedziału wieku, w którym zachorowania na badaną chorobę występują stosunkowo często. Takie założone z góry restrykcje na próbę dobieraną do badań ograniczają populację docelową i możliwość uogólniania wyników, ale nie stanowią źródła błędów analizy w tym konkretnym badaniu.

Wybór populacji osób eksponowanych

Grupa osób reprezentujących populację narażonych w badaniach kohort może pochodzić z rozmaitych źródeł. Wybór poszczególnej grupy zależy od różnorodnych względów: naukowych, ale także od dostępności badacza do tych osób lub ich dokumentacji i brana powinna być tu pod uwagę częstość występowania ekspozycji, potrzeba uzyskania kompletnej informacji wstępnej i danych z obserwacji w czasie prowadzenia badań. Zawsze odpowiedzi na te pytania winny być relatywizowane do stawianych hipotez, które badanie ma sprawdzić. Dla względnie częstych ekspozycji, jak palenie papierosów lub picie kawy, dostateczna liczba uczestników może być wybrana z wielu różnych populacji. Ale dla niektórych narażeń przemysłowych lub terenowych, właściwsze jest wybranie specyficznej populacji narażonych na rzadką ekspozycję, której

efekt zostanie poddany badaniu. Przykładami takich specyficznych ekspozycji może być praca z gumą, górnictwo uranu, sortowanie wełny, ale także specyficzna terapia jak np. naświetlanie promieniami X chorych z zeszywniającym zapaleniem stawów kręgosłupa. Kohortę narażonych mogą stanowić np. osoby żyjące w sferze środowiskowego narażenia, jak miejsca testów lub katastrof nuklearnych czy wysypiska odpadów toksycznych. Korzyścią płynącą z wyboru populacji specyficznego narażenia jest to, że daje on możliwość zgromadzenia dostatecznej próby do badań w rozsądnie krótkim czasie.

Użycie takich kohort może prowadzić do identyfikacji specyficznych czynników etiologicznych występujących w warunkach będących przedmiotem badania. Ponadto dostarczają one wydajnych metod identyfikacji czynników szkodliwych, występujących w populacji ogólnej. Jeżeli czynniki narażenia mają wpływ na powstawanie chorób, badanie ich efektów przy wyższych stężeniach może pozwolić na ustalenie związków przyczynowych w krótszym czasie obserwacji. Dotyczy to szczególnie tych substancji i innych narażeń, które mają charakter kumulacyjny.

Ponadto używanie kohort specyficznej ekspozycji pozwala na ewaluację rzadkich efektów (chorób), których analiza w innych warunkach wymagałaby nierealistycznie dużych kohort. Nawet jeśli dany efekt jest bardzo rzadki w populacji ogólnej, może on być wystarczająco częsty w próbie specjalnej ekspozycji, aby wywołać dostateczną liczbę przypadków w badaniu kohortowym. Na przykład roczna zachorowalność na *mesothelioma* w populacji ogólnej mieszkańców U.S.A. płci męskiej jest 8,4 na milion. Zatem prowadzona przez pięć lat kohorta 20 000 mężczyzn daje nikłą szansę uzyskania choćby jednego przypadku. Jednak w populacji robotników narażonych na pył azbestowy zachorowalność na *mesothelioma* jest dostatecznie częsta, aby kohorta 20 000 ludzi dała dostateczną liczbę przypadków do stwierdzenia wielu różnych zależności pomiędzy ekspozycją i chorobą. Mimo, iż badania kohort nie są optymalne w zastosowaniu do bardzo rzadkich chorób, to jeżeli procent różnicy ryzyka jest wysoki, badania kohort mogą być użyte efektywnie.

Ponieważ zasadniczym wymogiem z punktu widzenia ważności (*validity*) badania kohort jest uzyskanie kompletnej i dokładnej informacji o wszystkich

uczestnikach, szczególnie jeśli chodzi o dane dotyczące ekspozycji i efektu będącego przedmiotem badań, kohorty często są dobierane nie ze względu na specyficzną ekspozycję, ale ze względu na szanse zgromadzenia istotnych informacji. Na przykład badania wpływu codziennego przyjmowania małych dawek aspiryny na powstawanie chorób naczyniowych mózgu i serca były przeprowadzone na grupie lekarzy, którzy w słusznym przekonaniu autorów badania gwarantowali wysoką rzetelność podawanej informacji, ciągłe uczestnictwo w badaniach i regularne przyjmowanie leków. Inne podobnie dobierane grupy były rekrutowane spośród pielęgniarek, członków związków zawodowych, weteranów, absolwentów uczelni i według innych, podobnych kryteriów. Formowanie kohort z takich grup oferuje pewne logistyczne ułatwienia, jak istniejące gotowe listy adresów, ułatwiony dostęp do tych osób w czasie obserwacji, a także możliwość uzyskania gotowych danych informacji medycznej.

Ponieważ takie grupy nie są dobierane według ekspozycji, nadają się lepiej do badania efektów stosunkowo częstych. Założeniem takich badań jest również to, że grupy tak specyficznym dobrane mają reprezentować istotne z punktu widzenia badań cechy populacji ogólnej, aby wyniki uzyskane w tych badaniach można było uogólnić.

Wybór grupy ekspozycji do badania kohort powinien być dostosowany do postawionej hipotezy oraz specyficznych cech układu badań. Na przykład ocena możliwego związku narażenia na przemysłowy rozpuszczalnik z ryzykiem choroby nowotworowej pasuje najlepiej do retrospektywnej kohorty obejmującej szereg zakładów przemysłowych, używających tego rozpuszczalnika i posiadających rzetelnie prowadzoną dokumentację miejsc pracy i zadań poszczególnych pracowników. Natomiast w przypadku badania wpływu częstych ekspozycji na rozwój względnie częstej choroby przewlekłej najlepszym rozwiązaniem jest ogólna cohorta wylosowana z geograficznie lub demograficznie określonych miejsc, która może być analizowana pod kątem wyjściowych ekspozycji i następnie badana okresowo, aby wyłowić pojawiające się stany chorobowe. Klasycznym przykładem takiego typu kohorty jest wzmiankowane uprzednio badanie Framingham Heart Study. Fakt wyboru Framingham, nie był uzależniony od jakichś szczególnych charakterystyk ekspozycji, czy chorób w tym miejscu, ale zdecydował o tym zespół okoliczności, które pozwoliły badaczom wybrać i obserwować

uczestników przez wiele lat. Złożyła się na to stabilność populacji Framingham, różnorodność zawodów reprezentowanych przez mieszkańców miasteczka, istnienie dobrze prowadzonego pojedynczego szpitala obejmującego opieką jego mieszkańców i wreszcie bliskość Boston University, skąd pochodzą współtwórcy i realizatorzy projektu. Badania oparte na podobnym układzie kohort z populacji ogólnej przyczyniły się między innymi do wyjaśnienia przyczyn powstawania takich chorób jak choroba wieńcowa, choroba reumatyczna serca, zastoinowa niewydolność serca, chromanie przestankowe, dna moczanowa, choroby pęcherzyka żółciowego oraz licznych chorób oka.

Wybór grupy kontrolnej (referencyjnej)

Następnym krokiem po wyborze grupy osób ekspozowanych jest dobranie próby kontrolnej składającej się z osób nieekspozowanych. Jest to trudny i niesłychanie ważny krok w planowaniu badań ze względu na konieczność takiego doboru tej grupy, aby różniła się ona od grupy badanej tylko ekspozycją, a pod względem pozostałych cech możliwie jak najmniej. Różnice w doborze kohort mogą być źródłem poważnych błędów, niebezpiecznie obniżających konkluzyjność badań. Różnice w doborze grupy badanej i grupy kontrolnej mogą spowodować niewykrycie lub niedoszacowanie istniejących związków, albo nawet wykrycie związków tam, gdzie ich nie ma. Podstawową zasadą jest to, aby obie grupy pochodziły z tej samej populacji i aby poza badanym narażeniem różniły się jak najmniej między sobą.

Teoretycznie najlepszą grupę referencyjną (kontrolną) stanowiłyby te same osoby, które tworzą grupę narażonych, gdyby nie były narażone, ale jest to ideał nierealny. W praktyce w badaniach kohortowych dobierane są trzy typy grup kontrolnych:

- **Wewnętrzna grupa kontrolna** składa się z nienarażonych uczestników tej samej kohorty. W badaniu wpływu promieniowania rentgenowskiego w czasie fluoroskopowego prześwietlenia płuc na powstawanie raka sutka u kobiet chorych na gruźlicę, przykład takiej grupy mogą stanowić kobiety chore na gruźlicę, leczone w tym samym ośrodku medycznym, które nie były poddane tego typu naświetlaniom. Wewnętrzna grupa kontrolna, poza samym narażeniem, najmniej różni się od grupy badanej.

- **Populacja ogólna jako grupa kontrolna** sprawdza się lepiej w przypadkach, gdy grupa narażonych też jest dobierana z tej samej populacji. Bywa stosowana do porównania zapadalności lub umieralności empirycznej w określonych subpopulacjach z oczekiwaną zapadalnością większej populacji. W przypadku grup specyficznego narażenia stosowanie grupy kontrolnej z populacji ogólnej może być źródłem błędów spowodowanych tym, że grupy specyficznego ryzyka mogą się różnić od populacji ogólnej wieloma cechami, często trudno uchwytnymi dla badacza. Przykładami błędów wynikających z dobierania grupy kontrolnej z populacji ogólnej dla grup specyficznego narażenia jest tzw. efekt zdrowego pracownika oraz porównywanie grup pacjentów szpitalnych z osobami niehospitalizowanymi. Innym problemem może być trudny dostęp do danych osób tworzących populację ogólną i przez to pomijanie w analizie ważnych zmiennych.
- **Kohorta porównawcza** z innej, ale podobnej populacji, stanowi stosunkowo dobre odniesienia dla kohort specyficznego narażenia. W przypadku narażeń przemysłowych kohortę porównawczą mogą stanowić pracownicy innej fabryki lub nawet innego oddziału tej samej fabryki, w którym badane narażenie nie występuje. Potencjalny problem stanowi to, iż mogą występować znaczne różnice pod względem wielu innych niebadanych narażeń między pozornie podobnymi grupami pochodzącymi z różnych populacji.

W badaniach kohort, na początku których utworzona jest jedna ogólna kohorta, a następnie jej członkowie są zaklasyfikowywani do kategorii narażenia, może mieć zastosowanie wewnętrzna grupa referencyjna. Znaczy to, iż rezultaty uzyskane w grupie tych, którzy są poddani jakiejś szczególnej ekspozycji, są porównywane z rezultatami uzyskiwanymi u nieeksponowanych lub eksponowanych w innym stopniu. Na przykład badania kohorty lekarzy angielskich rozpoczęte przez Doll'a i Hill'a w 1950 roku dotyczyły porównania śmiertelności z powodu raka płuc osób, które nigdy nie paliły ze śmiertelnością z tego samego powodu osób palących różne liczby papierosów. Badacze stwierdzili znamienne różnice w umieralności z powodu raka płuc między palącymi i niepalącymi, a także stopniowy wzrost umieralności wraz ze wzrostem liczby wypalanych

papierosów. Podobnie w badaniach Framingham zostały oznaczone podstawowe poziomy cholesterolu oraz skurczowego i rozkurczowego ciśnienia tętniczego krwi, a następnie uczestnicy byli podzieleni na quantyle ekspozycji oparte na tych wartościach. W czasie obserwacji obliczano gęstości zachorowalności w tych grupach, a następnie porównywano je ze sobą. Należy zaznaczyć, że tam, gdzie wiele czynników ryzyka jest porównywanych jednocześnie, grupa nieeksponowanych winna być zdefiniowana jako ta, w której nie występuje żaden z czynników ryzyka poddanych ewaluacji. Zatem grupa kontrolna winna zawierać jednolitą, dobrze zdefiniowaną populację referencyjną, a nie jedną z kategorii czynnika narażenia, który jest w danej chwili analizowany.

Dla kohort, które używają specjalnej grupy ekspozycji, jak ma to miejsce w narażeniach zawodowych lub w szczególnych warunkach środowiskowych, niekiedy trudno jest wydzielić część kohorty wolną od badanego narażenia, aby użyć jej jako grupy referencyjnej. W takich przypadkach grupa kontrolna jest rekrutowana z populacji ogólnej w obszarze, gdzie zamieszkują ludzie należący do grupy narażonych lub nawet przyjmowana bywa jako grupa kontrolna, ogólna populacja danego kraju w tej samej grupie wiekowej jak populacja badana. Częstość występowania chorób jest wtedy obliczana w grupie badanej i odnoszona do oczekiwanej częstości występowania w populacji ogólnej w tym samym czasie obserwacji. Na przykład, aby oszacować potencjalne ryzyko pracy w przemyśle gumowym, umieralność w fabryce opon w Akron w stanie Ohio, była porównana z umieralnością w populacji ogólnej U.S.A. w tych samych grupach wieku. Badacze stwierdzili, że umieralność ze wszystkich powodów wynosi u pracowników fabryki opon zaledwie 82% umieralności wyliczonej dla całego społeczeństwa amerykańskiego.

Porównania takie są możliwe oczywiście tylko dla takich wskaźników, dla jakich istnieją dane ogólnokrajowe, na przykład dla umieralności, zachorowalności na nowotwory lub inne choroby poddane krajowemu nadzorowi epidemiologicznemu. Populacja ogólna w sposób oczywisty zawiera wszystkie populacje o specyficznych narażeniach. Dlatego taki sposób tworzenia ogólnokrajowej kontroli zakłada niską zachorowalność ogólną na badaną chorobę i małe rozpowszechnienie owych

specyficznych narażeń, aby błąd niedoszacowania wynikły z włączenia do grupy kontrolnej grup ludzi narażonych mogli być stosunkowo małe (<0.1).

Największym jednak niedostatkiem używania populacji ogólnej jako grupy referencyjnej jest to, że jej członkowie mogą nie być porównywalni bezpośrednio z kohortą badaną. Po pierwsze ludzie zatrudnieni są zwykle zdrowsi od tych, którzy nie pracują. Ponieważ ogólna populacja zawiera ludzi, którzy nie pracują z powodu słabego zdrowia na równi z pracującymi, grupy ludzi zatrudnionych mają zwykle niższe wskaźniki umieralności niż populacja ogólna. Jest to tak zwany w epidemiologii efekt zdrowego pracownika (*healthy worker effect*). Nawet te populacje pracownicze, u których narażenie na czynniki szkodliwe zwiększa szanse choroby i śmierci w porównaniu z pracownikami o podobnym charakterze pracy, ale nienarażonymi, mogą mieć ciągle niższą umieralność niż populacja ogólna. Ale nawet wtedy, gdy jako populacje porównawcze dobieramy grupy, które wydają się nam bardzo podobne do grupy badanej pod względem cech środowiskowych i demograficznych, mogą się one różnić od siebie bardzo pod względem takich, istotnych dla zdrowia cech, jak ilość wypalanych papierosów, spożycie alkoholu, czy poziom wysiłku fizycznego. Takie dane nie są znane dla całej populacji i mogą być źródłem błędów wywołanych czynnikami zakłócającymi, których nie można kontrolować z racji nieznaności ich charakteru i rozkładu.

Aby uniknąć tych problemów, zamiast stosować odniesienie do populacji ogólnej, używane są grupy referencyjne o podobnym przekroju demograficznym jak badana kohorta, ale składające się z osób nienarażonych. W badaniach ekspozycji przemysłowych takie porównywane grupy składałyby się z pracowników wykonujących różne typy pracy w różnych miejscach tego samego zakładu pracy, jedne w miejscach z narażeniami, a inne bez tych narażeń. Jeżeli są trudności z wydzieleniem w danej fabryce obszarów bez narażenia, można wydzielić grupy o różnej ekspozycji na zasadzie rodzaju wykonywanej pracy, niekoniecznie od razu na wstępie oceniając stopień narażenia. Po skalkulowaniu zachorowalności (czy umieralności) dla każdej z grup, uzyskane wyniki można odnieść do danych wyliczonych dla wszystkich uczestników badań. Tak uczyniono w badaniach nad narażeniem na kontakt z produktami używanymi w przemyśle gumowym w cytowanych już badaniach z Akron w Ohio, U.S.A. W ten sposób

uzyskujemy obraz efektów mniejszych lub większych od średniej dla całego zakładu, co pozwala dokonać mapowania zagrożeń w fabryce. W badaniach nad efektami defoliantów używanych w Wietnamie, grupę kontrolną stanowili podobni pod względem wieku, wyjściowego stanu zdrowia piloci wojskowi odbywający przeloty transportowe. W badaniach nad wpływem przewlekłej ekspozycji na promieniowanie rentgenowskie radiolodzy byli porównywani z internistami, którzy dzielili z nimi większość cech socjoekonomicznych i środowiskowych, za wyjątkiem ekspozycji na promieniowanie. W jeszcze innym badaniu pracownicy fabryki tkanin azbestowych zostali porównani pod kątem zachorowalności na raka płuc z pracownikami fabryki tkanin bawełnianych. Są to wzory starannego dobierania kohort, tak aby zminimalizować inne różnice, niż podstawowa różnica ekspozycji.

Bywa, że znalezienie grupy referencyjnej dostatecznie podobnej (poza ekspozycją) do grupy badanej jest trudne. Pewnym rozwiązaniem jest wtedy dobór kilku różnych grup referencyjnych. Jeżeli wszystkie one, różniąc się od siebie pod najrozmaitszymi względami dadzą podobne wyniki, tak samo odmienne od grupy badanej możemy wnosić, że różnice pomiędzy nimi są nieistotne z punktu widzenia badanej zależności i uznać uzyskaną zależność jako prawdopodobną. Jednak zawsze należy bardzo starannie rozważyć, czy nie mamy w naszych badaniach do czynienia z czynnikiem zakłócającym. Ponieważ taki czynnik (obyczajnie seksualne, status socjoekonomiczny, grupa etniczna) mógł występować przy porównaniu kobiet używających doustnych środków antykoncepcyjnych z kobietami, które nie używały żadnej antykoncepcji, i dlatego porównania dokonano z grupą kobiet stosujących środki antykoncepcyjne inne niż doustne.

Dobór losowy i dopasowywanie

Epidemiolog, który chce zbadać związek jednego określonego narażenia z efektem musi zdawać sobie sprawę z mnogości innych czynników, które mogą wpływać na uzyskany wynik, a które bardzo trudno jest kontrolować, a w wielu przypadkach trudno nawet uświadomić sobie ich obecność. Jeśli istnieją uzasadnione podejrzenia, iż jakiś czynnik ma wyraźny efekt zakłócający badanie związków statystycznych, metodą eliminacji jego wpływu jest stratyfikacja, podział kohorty na dwie mniejsze. Jedną z nich

będzie zawierała jednostki posiadające daną cechę, a druga nieposiadające jej. Jednak powtarzanie takiego podziału dla kolejnych zmiennych redukuje liczebność badanej próby w postępie geometrycznym i nawet po kilku powtórzeniach liczebność próby może nie wystarczyć do dalszej analizy. Uznana metoda redukcji wpływu czynników zakłócających jest dobór losowy, czyli taki, w którym każda jednostka w puli z której losujemy (*sampling frame*) ma jednakową szansę, że zostanie wylosowana. Losowanie nie eliminuje czynników zakłócających, ani innych konkurencyjnych czynników ryzyka, ale rozkłada je równomiernie w obu badanych grupach. Z taką sytuacją mamy do czynienia w badaniach eksperymentalnych, gdzie wybieramy grupę kontrolną i badaną jak najbardziej do siebie podobne, a następnie na jedną z tych grup nakładamy ekspozycję np. podajemy badany lek. W kohortach obserwacyjnych nie można zastosować doboru losowego do pełnego ujednoczenia rozkładu cech w grupach narażonych i nienarażonych, ponieważ każda z tych grup rekrutowana jest z innej subpopulacji. Rola losowania w tych badaniach polega na niestronniczym doborze uczestników badania z łącznej puli narażonych i nienarażonych, zanim dokonamy podziału na te kategorie, lub na niestronniczym doborze uczestników badania przy losowaniu osobno z puli narażonych i nienarażonych, jeżeli ten podział nastąpił przed rekrutacją uczestników badania. W obu przypadkach czynnik zakłócający powiązany z ekspozycją nie zostaje rozłożony równomiernie między kategoriami narażonych i nienarażonych.

Innym sposobem eliminacji lub zmniejszenia wpływu czynników zakłócających jest dobór dopasowany (*matching*). Dopasowanie ma na celu zapewnienie, że dwie grupy są zgodne pod względem czynników zakłócających.

Może ono mieć charakter zgodności indywidualnej w postaci ustawienia uczestników badania w pary o charakterystyce zgodnej ze względu na zmienną dopasowującą: palący z palącym, niepalący z niepalącym. Jest to dobór indywidualny (*pair matching*) lub dobór grup o tym samym odsetku czynnika dopasowującego, dobór proporcjonalny (*frequency matching*). Np. w badaniu wpływu pracy z azbestem na powstawanie raka płuc kohorta narażonych, pracujących z azbestem, miałaby tę samą

proporcję osób palących, co grupa kontrolna pracujących w podobnych warunkach z niezawierającą azbestu watą mineralną.

W badaniach kohortowych dopasowanie jest rzadziej stosowane niż w badaniach kliniczno-kontrolnych. Jednak dopasowanie w badaniach kohortowych stosunkowo dobrze eliminuje błędy wynikające z czynników zakłócających, bez żadnych specjalnych metod statystycznych. Może jednak osłabiać siłę związków statystycznych. Dopasowywanie ze względu na czynniki niemające wpływu zakłócającego związku między narażeniem a efektem zwykle osłabia siłę tych związków, nie dając żadnych innych korzyści.

Źródła danych w badaniach kohort

W projektowaniu badania kohort istotnym warunkiem wstępnym jest zapewnienie sobie metody uzyskania dokładnej i pełnej informacji, która pozwoli zaklasyfikować osoby uczestniczące w badaniach pod względem ekspozycji lub jej braku a także ocenić, które z osób włączonych do kohorty rozwinęły chorobę, czy inny badany wynik. Informacja dotycząca ekspozycji może być uzyskana z rozmaitych źródeł, w tym także z dokumentacji zgromadzonej niezależnie od prowadzonych aktualnie badań. Może pochodzić z dokumentacji lekarskiej, czy danych osobowych, informacji podanych przez same osoby badane w wywiadach lub kwestionariuszach; mogą to być dane pochodzące z badania lekarskiego uczestników lub z testów innego rodzaju, a także z bezpośrednich pomiarów środowiska, w którym członkowie kohorty żyją lub pracują. Dane o zachorowaniach, czy innych zmiennych wynikowych również mogą być uzyskane z istniejącej dokumentacji, w tym świadectw zgonów lub historii chorób, kwestionariuszy lub badania lekarskiego. Każde z tych źródeł ma swe wady i zalety, które należy brać pod uwagę w planowaniu, a potem w interpretacji poszczególnych badań.

Źródła informacji o ekspozycji

Metody i techniki używane do określenia ekspozycji w badaniach kohort zmieniają się zależnie od charakteru czynników ryzyka. Dla wielu kohort, dokumentacja istniejąca wcześniej, jak na przykład karty ambulatoryjne lub szpitalne historie choroby albo dokumentacja zatrudnionych prowadzona przez pracodawców mogą zawierać dane dostateczne do sklasyfikowania osób w zależności od ich statusu narażenia, a ponadto

dostarczyć koniecznych danych demograficznych lub innych. W niektórych przypadkach może to być jedyne dostępne źródło tego rodzaju informacji. Na przykład w kohorcie pacjentów poddanych naświetlaniom z powodu zeszywniającego zapalenia stawów kręgosłupa można z danych wywiadu dowiedzieć się o fakcie naświetlania, ale po informacje o wielkościach dawek należy sięgnąć do dokumentacji lekarskiej.

Korzystanie z już istniejącej dokumentacji ma wiele praktycznych zalet. Po pierwsze taka informacja zazwyczaj może być uzyskana dla znacznej liczby uczestników kohorty, a jej zgromadzenie nie wymaga dużych nakładów finansowych. Zatem, aby zapewnić dostateczne dane odnośnie ekspozycji będącej przedmiotem zainteresowania, słuszne jest zebranie możliwie najobszerniejszej informacji z istniejącej dokumentacji, nawet jeśli konieczne będzie korzystanie również z innych dodatkowych źródeł. Co więcej, dane z dokumentacji bywają gromadzone wcześniej niż zaplanowano dane badanie i przez osoby niezwiązane z nim, a zatem mają ważny walor bezstronności w ocenie statusu ekspozycji.

Dla wielu narażeń dokładność danych zawartych w istniejącej wcześniej dokumentacji może być niedostateczna, aby odnieść się adekwatnie do zasadniczego pytania badań. Ponadto taka dokumentacja zazwyczaj nie zawiera informacji odnośnie potencjalnych zmiennych zakłócających, takich jak obyczaje żywieniowe, palenie tytoniu, aktywność ruchowa oraz inne czynniki związane ze stylem życia. Te informacje mogą być dostarczone jedynie przez same osoby uczestniczące w kohorcie. Zatem wywiady i kwestionariusze wypełniane przez osoby badane są szczególnie ważne dla zebrania tych danych na temat ekspozycji, które nie są gromadzone w rutynowej dokumentacji. Potencjalny element stronniczości zawsze istnieje przy gromadzeniu tych danych bezpośrednio od uczestników już w trakcie prowadzonych badań. Uczestnicy badań, którzy są chorzy, lub mają swe uprzedzenia odnośnie związków będących przedmiotem badań, zwykle mają skłonność do udzielania odpowiedzi, których w ich mniemaniu oczekuje badacz. Z drugiej strony stygmatyzacja społeczna niektórych ekspozycji, takich jak picie alkoholu czy używanie narkotyków, może również wpływać na szczerść udzielanych odpowiedzi. W takich przypadkach, jeśli nie jesteśmy w stanie uzyskać

obiektywnej weryfikacji uzyskanych odpowiedzi, winniśmy się przynajmniej starać, aby dane były uzyskiwane w porównywalny i podobny sposób od wszystkich uczestników.

O niektórych ekspozycjach lub cechach zmiennych w czasie, takich jak ciśnienie tętnicze krwi lub poziom cholesterolu, dane zbierane w przeszłości nie dostarczają adekwatnej informacji, więc bezpośrednie badanie fizykalne lub testowanie krwi może być niezbędne. Dane, które zostaną uzyskane w sposób porównywalny dla wszystkich uczestników kohorty, mogą stanowić obiektywną i bezstronną podstawę klasyfikacji na ekspozowanych i nieekspozowanych.

Niemal nigdy uczestnicy badań nie znają swych poziomów narażenia na czynniki środowiskowe, jak szkodliwe substancje przemysłowe, czy zanieczyszczenia powietrza. W niektórych przypadkach można tu zastosować ekspozycje zastępcze, jak rodzaj wykonywanej pracy, czy stanowisko pracy w zakładzie, czy odległość od źródła zatrucia. Są to jednak w najlepszym razie tylko przybliżenia obarczone trudnym nieraz do oceny błędem. Ponadto dla wielu środowiskowych i zawodowych ekspozycji tego rodzaju sposób ich oceny nie może być stosowany. W takich wypadkach konieczne są bezpośrednie pomiary zanieczyszczeń powietrza czy wody w odpowiednio zdefiniowanych miejscach. Na przykład w wykonanym przez Uniwersytet Harvarda badaniu zanieczyszczenia powietrza w sześciu miastach, powietrze w badanych miejscowościach było wielokrotnie analizowane pod kątem poziomu dwutlenku siarki, masy wdychanych cząsteczek pyłu, siarczków, dwutlenku azotu i ozonu. Monitorowanie było powtarzane co sześć dni wewnątrz i na zewnątrz domów, a ponadto wyselekcjonowani ludzie byli monitorowani przez 24 godziny, odnośnie ich miejsc pobytu i zajęć, aby móc ocenić ich narażenie wobec niejednorodności stężeń polutantów w różnych miejscach miast. Taki typ monitorowania nie może mieć zastosowania do kohort retrospektywnych, które siłą rzeczy muszą opierać się na istniejącej dokumentacji. W ewaluacji ekspozycji przemysłowych mających miejsce w przeszłości ważna jest znajomość danych technologicznych i dat wprowadzania przepisów dopuszczających lub zakazujących stosowania interesujących nas substancji lub technologii. Niekiedy trzeba się wtedy odwołać do współpracy fachowców z danej dziedziny. Na przykład betanaftylamina mogąca być przyczyną raka pęcherza została wyeliminowana z

zastosowań w przemyśle gumowym w USA w 1949 roku, zatem pracownicy mogli być narażeni na kontakt z nią, tylko jeśli pracowali w tym przemyśle przed 1949 rokiem.

W wielu przypadkach zapewnienie odpowiedniej informacji na temat ekspozycji wymaga korzystania z wielu źródeł. Na przykład dane na temat różnych czynników ryzyka chorób serca w badaniach Framingham były uzyskiwane z kwestionariuszy odnośnie takich czynników jak wiek, palenie papierosów i wywiad rodzinny; z badań laboratoryjnych: poziomy w surowicy cholesterolu, cukru i trójglicerydów; badania fizykalnego: ciśnienie krwi, wskaźnik masy ciała, wzrost; oraz specjalnych badań: elektrokardiogram, badanie radiologiczne.

Bywa, że w badaniach kohort jednorazowe zaklasyfikowanie jest dokonywane w chwili włączenia do badań. Nie zawsze jednak ekspozycja pozostaje niezmienna w czasie badania. Ludzie mogą zmieniać liczbę wypalanych papierosów, stosowaną dietę, miejsce pracy. Zmiany technologiczne wprowadzane w zakładzie również mogą zmieniać poziom ekspozycji. Jeżeli wielkość ekspozycji ulegnie zmniejszeniu w czasie obserwacji kohorty, możemy otrzymać niedoszacowanie badanych asocjacji. W takich przypadkach wskazane jest regularne sprawdzanie ekspozycji w czasie obserwacji kohorty. Uważny czytelnik mógł spostrzec, że kohorta o zmiennym stopniu ekspozycji nie powinna być analizowana jako kohorta statyczna, lecz jako kohorta dynamiczna o wielu kategoriach ekspozycji. W tak prowadzonej kohorcie możemy nawet dołączyć zupełnie nowe ekspozycje już w czasie obserwacji. W badaniach Framingham po ośmiu latach prowadzenia kohorty, narastające zainteresowanie wpływem wysiłku fizycznego na powstawanie chorób serca spowodowało włączenie jako poziomu ekspozycji kategorii wydatku energetycznego.

Źródła danych zmiennej wynikowej (*outcome*)

Warunkiem wstępnym badania kohortowego jest uzyskanie kompletnej, porównywalnej i bezstronnej informacji na temat zdrowia każdej z obserwowanych osób w okresie prowadzenia kohorty. Ideałem byłoby wykonanie wszystkim jej uczestnikom specjalnych badań przez odpowiednio przeszkolony personel medyczny według jednolitej metodologii. Jednak taka sytuacja nie zawsze jest możliwa do zrealizowania. Stanowczo wymagana jest ona w kohortach szpitalnych wykorzystywanych w badaniach klinicznych oraz w bardzo zaawansowanych i

kosztownych kohortach terenowych jak np. *Framingham Heart Study*". W wielu badaniach źródła danych odnoszących się do zmiennej wynikowej (*outcome*) stanowi dostępna dokumentacja o ograniczonej i czasem niejednorodnej wiarygodności. Dlatego wstępna ocena spodziewanej jakości dostępnej informacji na temat uzyskiwanych rozpoznań jest warunkiem wstępnym do decyzji, czy podejmowanie danego badania rokuje wynik pozwalający na uzasadnione wnioski. Potencjalnymi źródłami informacji mogą być dane rutynowego nadzoru epidemiologicznego, świadectwa zgonów lub okresowe badania lekarskie. Niekiedy, aby uzyskać kompletne dane musimy się uciekać do kompilacji różnych źródeł. Dla chorób, które często kończą się śmiercią, można za daną wynikową przyjąć śmierć i informacji szukać w świadectwach zgonu. Są one stosunkowo łatwo dostępne, ale ich wiarygodność zależy od tego, co przyjmujemy za zmienną wynikową. Jeśli badamy ogólną umieralność, wiarygodność ta jest bardzo duża. Rzadko nie zauważa się śmierci i rzadko rozpoznaje się ją u żywych. Ale tam, gdzie chodzi o umieralność ze ściśle zdefiniowanej przyczyny, wiarygodność ta jest mniejsza i różna dla różnych chorób. Tylko w niewielkim procencie przypadków rozpoznanie przyczyny zgonu jest oparte na badaniu pośmiertnym, a w przeważającej liczbie przypadków na opinii lekarza domowego, który mógł nawet nie widzieć danej osoby przed jej śmiercią i opiera się na relacjach rodziny lub dawniejszych historiach chorób. Ponadto w chwili śmierci człowiek może mieć więcej niż jedną rozpoznaną chorobę i wybór głównej przyczyny zgonu nie jest łatwy. A jeżeli ponadto w kohorcie posługujemy się rozpoznaniem przyczyn zgonu o różnej wiarygodności u różnych osób, mamy do czynienia z potencjalnym źródłem poważnych błędów. Dla wyników innych niż śmierć dane mogą być uzyskane z notatek lekarskich, szpitalnych kart wypisowych, rejestrów chorób występujących w danej populacji oraz danych z kas chorych (czy innych systemów administrowania służbą zdrowia). Dane o chorobie mogą być również uzyskane bezpośrednio od uczestników w sposób, który ma podobne zalety i wady, jak przy zbieraniu danych o ekspozycji. Jeżeli uczestnicy badań są świadomi przedmiotu badań i stawianych hipotez, wzrasta ryzyko stronniczości danych uzyskiwanych w kwestionariuszach. Aby tego uniknąć, kwestionariusze bywają używane do identyfikacji tych osób, których dotyczy raport o zachorowaniu, ale po ich zidentyfikowaniu

uzyskiwane są dokładniejsze i bardziej bezstronne informacje mogące potwierdzić rozpoznanie. Te pochodzą ze szpitalnych historii chorób, czy raportów histopatologicznych. Dokładność diagnoz podawanych przez pacjentów zależy od rodzaju choroby. W amerykańskich badaniach nad zdrowiem pielęgniarek stwierdzono, że odnośnie chorób o jasno sprecyzowanych kryteriach diagnostycznych, takich jak złamania, nadciśnienie tetnicze, nowotwory piersi, skóry, jelita grubego i tarczycy ponad 90% diagnoz podawanych przez uczestniczki było potwierdzanych po sprawdzeniu dokumentacji medycznej. Ale dla innych stanów, takich jak nowotwory płuc, jajnika i macicy, dla zawałów serca i udaru mózgu potwierdzany bywa znacznie mniejszy procent rozpoznań uzyskanych z ankiet. Zatem dla wielu chorób dodatkowe potwierdzenia są konieczne. Istnieją również takie stany i choroby, w stosunku do których jedynym dokładnym i wiarygodnym źródłem informacji może być tylko okresowe badanie lekarskie i laboratoryjne. Podejście takie jakkolwiek znacznie bardziej kosztowne i czasochłonne, daje stosunkowo najbardziej obiektywne i wiarygodne źródło informacji o chorobie i pozwala na standaryzację używanych procedur diagnostycznych zwiększając powtarzalność i porównywalność wyników. Wiele badań epidemiologicznych chorób serca z Framingham włącznie, opiera się na okresowych badaniach. Ponieważ kompletne badanie lekarskie zawiera w sobie również wywiad, jest ważne, aby osoba przeprowadzająca badanie nie znała stanu ekspozycji osoby badanej. Stosowanie tych samych kryteriów diagnostycznych w obu porównywanych grupach jest kluczowe dla wiarygodności wyników. Większość epidemiologów woli polegać na mniej dokładnych, ale ujednoliconych kryteriach niż na wybiórczej poprawie dokładności przez diagnozowanie jednej grupy w szpitalu, a drugiej w rejonie. Typowym przykładem jest porównanie zachorowań grupy osób poddanych ekspozycji w przemyśle, które są regularnie badane przez lekarzy z zachorowalnością na tę samą chorobę w populacji ogólnej, gdzie wiele rozpoznań może być przeoczonych. Stronnicza ocena zagrożenia u eksponowanych może tu wynikać z lepszego znajdowania osób chorych, a nie z rzeczywiście częstszych zachorowań. Jeżeli więc rozpoznania przyczyn śmierci w populacji ogólnej opieramy na świadectwach zgonu, należy w grupie osób narażonych też

oprzeć się na świadectwach zgonu, nawet jeśli wszystkim narażonym, którzy zmarli, wykonano po śmierci autopsje.

Badania epidemiologiczne nie ograniczają się do analizy zachorowań. W niektórych z nich efektami badanymi mogą być inne zmienne będące przedmiotem zainteresowania w dziedzinie zdrowia publicznego. Jako przykłady można podać efekty promocji zdrowia np. zmiany diety, rezygnacja z używek, czy inne zachowania prozdrowotne. W takich przypadkach materiałem będącym podstawą analizy mogą być odpowiednio przeprowadzone ankiety lub autoraporty dotyczące trybu życia, sposobów spędzania czasu i zjadanych potraw.

Problemy związane z przebiegiem obserwacji

W każdej kohorcie, czy to prospektywnej, czy retrospektywnej uzyskanie danych odnośnie wystąpienia efektu (*outcome variable*) wymaga śledzenia losu uczestników od momentu zakwalifikowania do grupy eksponowanych lub nieeksponowanych do momentu zakończenia obserwacji, aby określić, czy i kiedy badany efekt wystąpił. Niemożność uzyskania tej informacji dla każdego z uczestników, a szczególnie różnice w liczbie osób, które wypadły z obserwacji między grupą eksponowanych i nieeksponowanych jest potencjalnym źródłem stronniczości i może sprawić, że badanie okaże się bezwartościowe. Prowidłowe prześledzenie kohorty, a czasem wymaga to wielu lat starannego gromadzenia informacji, jest zawsze poważnym wyzwaniem dla badaczy, a także powodem wysokich kosztów tego typu badań.

Czas niezbędny do prawidłowego przeprowadzenia badania kohorty jest zależny od czasu indukcji (inkubacji) choroby. Może on wynosić dni lub tygodnie jak w wypadku niektórych ostrych chorób; miesiące w przypadku wad wrodzonych lub poronień spontanicznych, a nawet lata, gdy mamy do czynienia z chorobami o charakterze przewlekłym jak nowotwory, czy niektóre choroby serca. Kohorty wymagające dłuższej obserwacji są bardziej narażone na utraty uczestników z powodu zmiany miejsca zamieszkania lub miejsca pracy, czy po prostu rezygnacji z udziału w badaniach. Czasem przy dużym wysiłku prowadzących badania udaje się zmniejszyć liczbę utraconych z badania osób przez wykorzystanie innych, pośrednich źródeł informacji o ich stanie zdrowia, jak rejestry chorób, dokumentacja szpitalna i sanatoryjna, czy akty zgonów lub

dokumentacja przedsiębiorstw ubezpieczeniowych. Zawsze jest to jednak bardzo duży wysiłek i poważne zwiększenie czasu pracy.

Nie we wszystkich przypadkach wynik badany w badaniach kohortowych stanowi wystąpienie badanej choroby lub zgonu z jej powodu. W wielu przypadkach zmienną mierzoną jest czas upływający do wystąpienia danego zdarzenia: zgonu, zachorowania, czy zdefiniowanego zespołu objawów. Taka analiza mająca np. postać krzywych Kaplana - Meiera jest bardzo ciałym narzędziem porównywania efektów badanych narażeń.

Analiza danych w badaniach kohort

W prospektywnym badaniu kohortowym dobór uczestników badania może być dokonywany na jeden z dwóch sposobów:

- 1) W badanej populacji lub próbie składającej się z osób wolnych od badanego efektu, np. choroby sprawdzamy stan narażenia uczestników i dzielimy naszą próbę na dwie kohorty: narażonych i nienarażonych.
- 2) Wolne od badanego efektu osoby narażone i nienarażone są dobierane indywidualnie i włączane do badania.

W kohortach stacjonarnych dobór narażonych (eksponowanych) i nienarażonych dokonywany jest na początku badania i tak utworzony zbiór jest przedmiotem obserwacji przez założony z góry okres czasu i w każdej z wydzielonych grup rejestrowane są przypadki zachorowań lub występowanie innego badanego efektu, a po zakończeniu badania określana jest zapadalność wśród narażonych i nienarażonych. Porównanie tych zapadalności dokonywane jest przez wyliczenie ich stosunku, który nazywamy ryzykiem względnym (*relative risk-RR*) oraz ich różnicy nazywanej też ryzykiem przypisanym.

W wielu przypadkach, zwłaszcza tych, w których mamy do czynienia z rzadko występującym narażeniem, taki dobór byłby bardzo trudny, dlatego osoby narażone są włączane do badania w miarę jak zostają zidentyfikowane, a osoby nienarażone są dobierane do nich na zasadzie doboru losowego lub dopasowanego. W tak zaplanowanym badaniu poszczególni uczestnicy mogą być obserwowani przez różne okresy czasu, a zatem nie ma możliwości obliczenia średniej zapadalności dla kohort. Wobec czego stosowana jest miara nazywana gęstością zapadalności, która stanowi stosunek liczby osób, które w danej kohorcie zachorowały do sumy osobo-czasu

obserwacji zarówno tych osób, które zachorowały jak tych, które nie zachorowały. Ta miara osobo-czasu (np. osobo-lat, osobo-dni) stanowi sumę czasu obserwacji wszystkich uczestników danej kohorty mierzonego osobno dla każdego uczestnika.

W badaniach kohortowych zarówno określenie stanu narażenia, jak i choroby musi dotyczyć każdego indywidualnego uczestnika badania. Ale nie zawsze uczestnicy badania są do próby dobierani indywidualnie. W pewnej odmianie badań kohortowych można dobierać przez losowanie całe grupy ludzi np. klasy szkolne, pracowników instytucji, mieszkańców pojedynczych bloków. Dobór taki nosi nazwę doboru grupowego (*cluster sampling*). Bywa on bardziej dogodny od doboru indywidualnego, ale stawia wyższe wymagania analizie statystycznej. Aby uzyskać znamienność statystyczną w takich badaniach potrzebne są mniejsze wartości wskaźnika „p” oraz węższe przedziały ufności. Pamiętać należy, że dobór grupowy nie zwalnia z obowiązku przypisania narażenia i efektu każdemu indywidualnemu uczestnikowi badania¹.

Weźmy prosty, fikcyjny przykład badania kohort. W wytwórni odczynników chemicznych kierownik działu osobowego spostrzegł, że pracownicy jednego z wydziałów częściej niż w innych oddziałach otrzymują zwolnienia lekarskie z powodu infekcji dróg oddechowych i zgłosił ten fakt lekarzowi poradni zakładowej. Ten po przejrzeniu dokumentacji medycznej pracowników tego oddziału stwierdził, że według posiadanej dokumentacji spośród 200 osób tam zatrudnionych, dwanaście choruje na przewlekłe obturacyjne zapalenie oskrzeli. Jak przekonał się na podstawie danych epidemiologicznych kraju, jest to odsetek znacznie wyższy niż w ogólnej populacji w tej grupie wieku. Postanowił zbadać, czy warunki pracy w tym oddziale są odpowiedzialne za występowanie u pracowników przewlekłego zapalenia oskrzeli. Spośród pracowników tego oddziału, którzy nie mieli przewlekłego zapalenia oskrzeli wybrał losowo grupę 100 osób w przedziale wieku 35-55 lat, a następnie według tych samych kryteriów wieku i zdrowia wybrał 100 osób z innego oddziału, o podobnym natężeniu intensywności pracy, zapylenia i hałasu, z których również żaden z pracowników nie chorował na przewlekłe zapalenie oskrzeli. Pierwszą grupę oznaczył jako kohortę

¹ Elementarne wzory na wyliczenie ryzyka względnego i różnicy ryzyka w badaniach kohortowych są podane w słowniczku zamieszczonym w „Niezbędniku epidemiologa” dlatego tu rezygnujemy z ich zamieszczania.

eksponowanych „E”, a drugą jako nieeksponowanych „N”. Na tym etapie badań narażeniem była praca w rzeczonym oddziale. W czasie pięcioletniej obserwacji stwierdził wystąpienie przewlekłej obturacyjnej choroby płuc (POChP) u następujących osób:

PRACOWNIK	KOHORTA	DATA ROZPOZNANIA	LICZBA DNI OBSERWACJI	LICZBA LAT OBSERWACJI
	01.01.1990- 31.12.1994	POChP		
AB	E	05.03.1990	64	0.18
CD	E	02.05.1992	853	2.34
DE	E	26.07.1992	938	2.57
EF	E	08.08.1992	951	2.61
GH	N	11.07.1993	1288	3.53
HI	E	03.10.1993	1341	3.67
IJ	N	05.12.1993	1404	3.85
JK	E	12.09.1994	1685	4.62
LN	E	15.11.1994	1749	4.79

Przyjmijmy, że nie było ubytków osób uczestniczących w badaniach, ani z powodu zgonów, ani z innych powodów. W oddziale E stwierdzono 7 nowych przypadków przewlekłego zapalenia oskrzeli, a w oddziale N tylko dwa.

Stąd możemy obliczyć zachorowalność zbiorczą dla każdej z kohort:

$$CIE = \text{Liczba nowych przypadków choroby} / \text{liczba obserwowanych osób} = 7 / 100$$

$$CIN = \text{Liczba nowych przypadków choroby} / \text{liczba obserwowanych osób} = 2 / 100$$

a następnie ryzyko względne: $RR = CIE / CIN = 7 / 2 = 3.5$ oraz różnicę ryzyka na tysiąc osób na rok: $AR = CIE - CIN = 1.4 / 100 - 0.4 / 100 = 1 / 100 = 10 / 1000$. Praca w oddziale E powodowała w skali rocznej 10 zachorowań na przewlekłe zapalenie oskrzeli w przeliczeniu na tysiąc osób zatrudnionych w tym oddziale.

Traktując te kohorty jako kohorty dynamiczne możemy wyliczyć gęstość zachorowalności u osób z oddziału E:

$$IDE = 7 / (93*5 + 0.18 + 2.34 + 2.57 + 2.61 + 3.67 + 4.62 + 4.79) \text{osobo-lat} =$$

$$= 7 / 485.78 \text{osobo-lat} = 1.44 / 100 \text{osobo-lat}$$

$$IDN = 2 / (98*5 + 3.53 + 3.85) \text{osobo-lat} = 2 / 497.38 \text{osobo-lat} = 0.40 / 100 \text{osobo-lat}$$

Ryzyko względne obliczone na podstawie gęstości zachorowalności wyniosłoby:

$RR = 6.35 / 1.90 = 3.34$, co nie odbiega wiele od wyniku uzyskanego w kohorcie stacjonarnej.

Przykład ten pozwoli nam również sprawdzić dokładność wyliczenia zachorowalności zbiorczej z gęstości zachorowalności według formuły wykładniczej:

$$CI(t) = 1 - \exp(-ID * \Delta t)$$

$$CIE(5) = 1 - \exp(-0.0144 * 5) = 0.0695 = 6.95 / 100$$

$$CIN(5) = 1 - \exp(-0.004 * 5) = 0.0198 = 1.98 / 100$$

Miary związku, stosunek i różnica ryzyka, wyliczane w badaniach epidemiologicznych muszą być zawsze podawane z przedziałem ufności. Pakiety statystyczne podają te przedziały gotowe, ale epidemiolog powinien znać formuły na podstawie, których są wyliczane. Określanie znamienności statystycznej wyników nie jest wystarczające, choćby z tego powodu, że nie daje ono informacji i sile związku zmiennych, a tylko o tym, jakie jest prawdopodobieństwo, że między badanymi zmiennymi występuje związek statystyczny. Natomiast ryzyko względne podaje siłę związku między zmiennymi, a różnica ryzyka, a szczególnie procent różnicy ryzyka podaje jaka część badanego efektu jest zależna od danego narażenia.

Istotnym elementem projektowania badań epidemiologicznych jest określanie liczebności próby (*sample size estimation*). Podstawowe formuły tych kalkulacji zakładają określenie pożądaných wartości znamienności i mocy testu, które stanowią założenia badacza, ale również przewidywanych wartości badanego parametru (np. zapadalności) oraz jego wariancji dla populacji. W niektórych przypadkach przybliżone oszacowania tych wielkości można uzyskać na podstawie wcześniejszych badań lub badań pilotowych. Choć są to oszacowania orientacyjne, stanowią one ważny punkt wyjścia i narzędzie pomocne w racjonalizacji kosztów badania.

Sposób wyliczania liczebności próby musi być dostosowany do charakteru badania, a przede wszystkim do sposobu doboru próby. Dlatego osoby z ograniczoną znajomością statystyki powinny wybór formuły wyliczania konsultować ze statystykiem.

Nie zawsze przedmiotem analizy w badaniach kohortowych jest fakt wystąpienia badanego efektu. W niektórych przypadkach, zwłaszcza gdy wystąpienie efektu jest nieodwołalne lub ma bardzo wysoki stopień prawdopodobieństwa, właściwsze jest badanie czasu, jaki upływa do wystąpienia efektu. Tego rodzaju analiza w postaci tzw. krzywych czasu do zdarzenia (*time to event curve*) jest przedmiotem osobnego opracowania.

Podstawowe błędy stroniczości w projektowaniu i prowadzeniu badań kohortowych:

Błąd wypadania z obserwacji

Jeżeli obserwowane osoby wypadają z kohorty stacjonarnej, dokładne wyliczenie zachorowalności zbiorczej staje się niemożliwe. Jednym z wyjść jest wtedy potraktowanie tej kohorty jako kohorty dynamicznej i obliczenie gęstości zachorowalności w oparciu o możliwe do prześledzenia czasy obserwacji w sposób podany powyżej. Można też określić zakres możliwych wyników przyjmując dwa ekstrema. Jedno ekstremum uzyskamy przyjmując, iż wszystkie osoby, które wypadły z kohorty zachorowały, a drugie przyjmując, iż żadna z nich nie zachorowała. Wtedy możemy oszacować iż „prawdziwa” zachorowalność zbiorcza mieści się gdzieś w przedziale pomiędzy wyliczonymi w ten sposób minimum i maksimum. Jeżeli kohorta jest duża, a straty nieliczne, zakres możliwego błędu jest niewielki i może być w niektórych przypadkach pominięty. Ale przy niezbyt licznych kohortach i rzadkich chorobach nawet pojedyncze przypadki wypadnięcia z obserwacji mogą być źródłem poważnych błędów.

Błędy spowodowane wypadaniem poszczególnych osób z kohorty mogą być przypadkowe, losowo rozłożone w grupach osób eksponowanych i nieeksponowanych lub różnicowe, gdy wypadanie z kohorty ma inne prawdopodobieństwo w grupie narażonych niż w grupie nienarażonych na ekspozycję, albo jeżeli zależy od badanej choroby (czy innego wyniku). Niekiedy zależec może zarówno od ekspozycji, jak i od

choroby. W tych przypadkach zarówno wielkości błędu mogą być bardzo duże, a jego kierunek niemożliwy do przewidzenia. Może się na przykład zdarzyć, że w badaniach wpływu palenia papierosów na powstawanie raka płuc osoby palące, które zachorują na raka łatwiej porzucają kohortę niż osoby niepalące, które zachorują. W takim przypadku oszacowanie asocjacji pomiędzy paleniem a rakiem będzie stronnicze i nie odda rzeczywistej siły związku.

Błąd odmowy uczestnictwa

Ważnym potencjalnym źródłem błędu w badaniach kohort jest efekt nieuczestnictwa. W każdej, zwłaszcza szerzej zakrojonej kohorcie tylko część osób kwalifikujących się do uczestnictwa zgadza się wziąć udział w badaniach. Na przykład w badaniach Framingham wstępne uczestnictwo wynosiło 69% (4469 osób z 6507), a w badaniach zdrowia pielęgniarek (Nurse's Health Study) wyniosło ono 71%. Na 172 tysiące listów zapraszających do uczestnictwa w badaniach odpowiedziało około 122 tysiące osób. Rzadko też osoby odmawiające rozkładają się losowo wśród osób zapraszanych do uczestnictwa. Zwykle osoby, które odmawiają uczestnictwa w badaniach, różnią się od uczestników stosunkiem do swego zdrowia, trybem życia i często generalnie postawą życiową. Na przykład w przeprowadzonych na lekarzach badaniach Doll'a i Hill'a nad efektem palenia papierosów, palący odmawiali częściej uczestnictwa niż niepalący. W badaniach Framingham proporcjonalnie więcej osób, które odmówiły uczestnictwa, zmarło w ciągu pierwszych dwóch lat prowadzenia kohorty, niż miało to miejsce wśród uczestników badań. W takich przypadkach powstaje pytanie o słuszność uogólnień (*generalizability, external validity*) opartych na tych badaniach. Istnieje możliwość, że stwierdzony w badaniach związek nie występuje lub ma inne nasilenie wśród osób, które odmówiły uczestnictwa w badaniach. Nie ma reguł logicznych, które określałyby wartość uogólniającą badań. Rozstrzyga osąd i rozsądek badacza, bądź czytelnika, który decyduje, czy różnice pomiędzy grupą poddaną badaniom a populacją docelową badań pozwalają na uogólnienie badań, czy nie. Odmowa uczestnictwa nie wpływa jednak na siłę związków stwierdzanych (*internal validity*) w badaniach, o ile nieuczestnictwo nie jest związane z badaną ekspozycją i jednocześnie z jakimiś innymi czynnikami ryzyka badanej choroby. Jeżeli więc na przykład nałogowi

palacze mniej licznie uczestniczą w badaniach, badania niedoszacują liczbę palaczy w populacji ogólnej, ale wpływ palenia na powstawanie badanej choroby będzie nadal zbadany właściwie. Jednak jeśli odmowa uczestnictwa byłaby powiązana z nałogowym paleniem i z innym jeszcze czynnikiem ryzyka badanej choroby, efekt palenia może być oszacowany błędnie. Podobnie jeśli chodzi o wpływ choroby na nieuczestnictwo. Jeśliby tylko ludzie o słabszym zdrowiu odmawiali uczestnictwa w badaniach, badania niedoszacują występowania słabego zdrowia w populacji, ale związek pomiędzy paleniem i chorobą będzie przedstawiony właściwie. Gdyby jednak osoby odmawiające udziału w badaniach byli jednocześnie palaczami i niezależnie od palenia mieli inne niż pozostali uczestnicy ryzyko choroby, na przykład raka płuc, wtedy ocena zależności pomiędzy paleniem i rakiem płuc mogłaby stać się stronnicza (*biased*). Rzadko badacz posiada dostateczne dane, aby ocenić, czy odmowa uczestnictwa nie wiąże się z ekspozycją, z chorobą czy z innymi niż ekspozycja czynnikami ryzyka choroby. Niekiedy jednak można obliczyć konsekwencje założeń istnieniu takich związków. W badaniach Doll'a i Hill'a wniosek, że palenie sprzyja powstawaniu raka płuc został zakwestionowany przez statystyków reprezentujących kompanie tytoniowe na tej właśnie podstawie, że nałogowi palacze mniej chętnie uczestniczyli w badaniach. W odpowiedzi autorzy wyliczyli, że aby uzyskać pożądaną dla kompanii tytoniowych efekt braku związku pomiędzy paleniem i rakiem przy pełnym uczestnictwie w badaniach, palenie tytoniu u palaczy nieuczestniczących musiałoby mieć efekt ochronny zmniejszający trzydziestokrotnie szanse zachorowania na raka płuc. Niekiedy po to, aby ocenić słuszność uogólnień odwołujemy się do podobieństw demograficznych pomiędzy uczestnikami i osobami które odmówiły uczestnictwa w badaniach. Jeśli wiek, płeć i inne charakterystyki obu grup są podobne, możemy rozsądnie przyjąć podobieństwo relacji pomiędzy ekspozycją i chorobą w obu grupach.

Błąd przynależności

Wymóg, aby w prolektowaniu badania kohortowego hohorta narażonych i nie-narażonych, poza czynnikiem narażenia, różniły się od siebie jak najmniej natrafia na szczególne trudności, gdy narażenie jest związane z określonym typem działalności lub szczególną charakterystyką grupy społecznej czy zawodowej. Typowym przykładem

błędu w doborze kohorty kontrolnej (nienarażonych) jest opisywany wyżej efekt zdrowego przecownika. Taka szczególna charakterystyka jednej z kohort może skutkować zmianą wielkości stosunku ryzyka, a nawet jego odwróceniu, jak to miało miejsce w przytoczonych wyżej badaniach efektów zdrowotnych pracy w przemyśle gumowym. Błąd ten może być szczególnie duży, jeżeli efekt narażenia ma niską swoistość, czyli może być powiązany z innymi czynnikami poza badanym narażeniem a występującymi również w grupie kontrolnej. Takimi nieswoistymi wskaźnikami jest absencja chorobowa, ale też np. umieralność ogólna.

W projektowaniu badań ważne zatem jest, aby dobrze rozeznąć charakterystykę osób narażonych i warunki w jakich narażenie występuje. Robotnicy przemysłowi mogą być nie tylko ogólnie zdrowsi od osób w tym samym przedziale wiekowym z populacji ogólnej, ale mogą też być zwalniani lub przesuwani do innej pracy w przypadku wystąpienia u nich zaburzeń zdrowia, stwarzając w wyniku selekcji dodatkowe przesłanki dla lepszego zdrowia populacji narażonych.

Stronniczość rozpoznania choroby

Znajomość narażenia może wpływać na czułość i specyficzność rozpoznania choroby szczególnie w tych sytuacjach, gdy stosowane metody diagnostyczne są trudne jak ma to miejsce w przypadku rozpoznania niektórych nowotworów. Materiał od osób narażonych jest badany skrupulatnie z obawy o możliwość nierozpoznania groźnej choroby, natomiast materiał od osób nienarażonych ma małe prawdopodobieństwo wstępne rozpoznania i bywa badany powierzchownie.

Przy tych wszystkich zastrzeżeniach badania kohort stanowią szczególnie precyzyjny układ badawczy w badaniach obserwacyjnych i stanowią wzór dla układu badawczego prób klinicznych. W kohortach prospektywnych ekspozycja jest określana przed wystąpieniem choroby, a zatem znajomość choroby nie może wpływać na kwalifikowanie badanych, jako eksponowanych lub nieeksponowanych, co stanowi istotne źródło stronniczości w badaniach kliniczno-kontrolnych. Ponadto badania kohort nadają się szczególnie do badania rzadkich ekspozycji, a także pozwalają na badanie więcej niż jednego efektu danej ekspozycji. Pozwalają one analizować relacje czasowe pomiędzy ekspozycją i chorobą oraz wyliczać bezpośrednio zapadalność wśród osób

eksponowanych i wśród nieeksponowanych. Kohorty są jednak mało efektywne dla badania rzadkich chorób, z wyjątkiem sytuacji, kiedy procentowa różnica ryzyka jest duża. Podstawowym problemem w prospektywnych badaniach kohort jest ich wysoki koszt oraz wielka wrażliwość wyników na wypadanie uczestników z obserwacji.