

ARALEZ HEALTH, LLC

Warsztaty PATH

Poprawa jakości opieki zdrowotnej i podejmowanie decyzji strategicznych

Warszawa, 17 listopada 2011 r.

Vahé A. Kazandjian

Niniejszy pakiet stanowi podsumowanie kluczowych zagadnień, poruszonych w trakcie warsztatów. Zainteresowani mogą uzyskać dodatkowe informacje i stosowne piśmiennictwo.

Wskaźniki i ich praktyczne zastosowania

Podczas warsztatów omówimy i poddamy dyskusji wskaźniki jako narzędzia dla użytkowników. Powyższe stwierdzenie niesie dwie bezpośrednie konsekwencje:

1. Trzeba zrozumieć cel, jakiemu służy narzędzie; oraz
2. Użytkownikiem narzędzia jest ten, kto propaguje właściwe rezultaty wynikające z wykorzystania narzędzia.

Przyjrzyjmy się bliżej tym zagadnieniom.

A. Wskaźniki jako narzędzia

W zasadzie, rolą wskaźnika jest “wskazywać”, nie “interpretować”. A zatem, termometr może wskazywać ciepłotę ciała 40°C, ale nie mówi, czy to jest hipertermia, czy infekcja. Tak samo, liczba cięć cesarskich pokazuje jakąś wartość numeryczną, ale nie mówi czy decyzja o porodzie przez cięcie cesarskie była decyzją właściwą.

Dlatego w pierwszym etapie istotne jest by dysponować dobrze skalibrowanym narzędziem, które nieomylnie i za każdym razem spełnia swoje zadanie. To dziedzina projektowania i testowania wskaźników, która zajmuje się dwoma podstawowymi atrybutami narzędzia – poprawnością i wiarygodnością, zdefiniowanymi jako:

Validity – Istotność: cecha narzędzia/miary pozwalająca ilościowo lub jakościowo określić to, co zamierzamy zmierzyć. Oznacza to, że projektujący dane narzędzie powinni mieć bardzo precyzyjnie określony cel i wskazać użytkownikom do jakiego pomiaru służy określone narzędzie.

Reliability -Wiarygodność: możliwość danego narzędzia/miary do ilościowego lub jakościowego określenia tego, co zamierzamy zmierzyć, z minimalną zmiennością lub błędem. Jest to atrybut bardziej „elastyczny” niż istotność poprawność, ponieważ wiarygodność zawiera się w pewnym zakresie, a nie jest układem „tak lub nie”. Na przykład, kiedy mówimy o istotności, nie można powiedzieć, że miara jest „istotna w 50%” – miara jest istotna, albo nie. Można natomiast powiedzieć, że miara jest “wiarygodna w 50%” implikując, że błąd w pomiarze wynosi 50%.

W sposób oczywisty, dobre narzędzie lub miara powinna być w 100% istotna i wysoce wiarygodna.

B. Użytkownicy narzędzia

Wykorzystywanie określonych narzędzi obejmuje zawsze zastosowanie „sztuki i wiedzy”. W naszej sytuacji, skoncentrujemy się wyłącznie na „wiedzy” jako przedmiocie naszych rozważań, ponieważ wskaźnik jest jednolitym i domniemanym konstruktem, który dąży do wyeliminowania subiektywizmu w obserwacji danego zjawiska. Co to znaczy?

Subiektywny pomiar zawiera niejako „wkomponowaną” wartość. Na przykład, chirurg może powiedzieć: „Podawanie antybiotykowej profilaktyki okołoperacyjnej przez cztery dni jest prawidłowe, gdyż zawsze tak z powodzeniem postępowalem”.

Obiektywny i jednoznaczny pomiar opiera się na dowodach: przykładowo proponuje wytyczne dla stosowania antybiotykowej profilaktyki okołoperacyjnej z uwzględnieniem czasu podawania antybiotyku przed i po zabiegu, określonej charakterystyki pacjenta, schorzeń współistniejących, typu operacji itp.

A zatem użytkownik narzędzia powinien wiedzieć jak interpretować kwantyfikację miary bez wartości (subiektywności) „wbudowanej” w określony pomiar.

Jednakże sama wartość (value) jest istotna i zróżnicowana w zależności od określonej kultury czy też kraju. W dalszej kolejności rozważymy w jaki sposób można uwzględnić wartość przy interpretacji określonych wskaźników.

Interpretacja wskaźników – wytyczne i wartości

Interpretacja wskaźnika uwzględnia następujące parametry:

- wielkość
- zmianę w czasie
- populację (pacjentów), której dotyczy określony wskaźnik
- wiarygodność pomiaru

Rozpatrzmy każdy z nich, krok po kroku.

A. Wielkość

Wielkość wskaźnika jest tylko liczbą, dopóki się jej z czymś nie porówna. To „coś” nazywane jest „normą”, lub „punktem odniesienia”.

Norma to wcześniej ustalony punkt docelowy w zakresie prawidłowości. Na przykład, dla szczepień przeciwko polio, norma może wynosić 100%; normatywny czas podania aspiryny pacjentom z bólem w

okolicy mostka i podejrzeniem zawału serca może wynosić 40 minut; a normatywny czas włączenia antybiotykowej profilaktyki okołoperacyjnej przed zabiegiem chirurgicznym może wynosić „x” godzin.

Punkt odniesienia to względna wartość porównawcza. Może stanowić porównanie z poprzednią wartością u tego samego podmiotu, lub porównanie „z” lub „do” wartości jaką uzyskał inny podmiot. Na przykład, współczynnik zgonów z powodu posocznicy w waszym szpitalu w styczniu, można porównać ze współczynnikiem zgonów tej samej kategorii w styczniu poprzedniego roku, a także ze współczynnikiem zgonów w innym szpitalu w Polsce lub w innym kraju.

Norma obowiązuje we wszystkich przypadkach i każdego; punkt odniesienia jest elastyczny i określany dla poszczególnych sytuacji.

B. Zmiana w czasie

Określana jest jako zmienność, może być sezonowa, lub wywołana zmianą praktyki. W obu przypadkach, zrozumienie przyczyn zmienności stanowi istotę jej interpretacji.

Jednak przed każdą interpretacją zmienności należy ocenić jakość danych – zmienność może być bowiem spowodowana niedoskonałym, w tym niekompletnym procesem zbierania i transmisji danych.

C. Populacja/Pacjenci

Narzędzie jakie stosujemy musi być skorygowane, lub skalibrowane w odniesieniu do charakterystyki mierzonego zjawiska. W przypadku wskaźników opieki zdrowotnej, trzeba uwzględnić charakterystykę demograficzną i kliniczną pacjentów oraz korekty dotyczące kryteriów takich jak ciężkość określonego przypadku, case-mix (czyli LICZBĘ pacjentów o określonej charakterystyce w STOSUNKU DO ogólnej liczby leczonych chorych) i specyfikę poszczególnych schorzeń. We wszystkich przypadkach, oczekuje się, że za pomocą wskaźnika dokonamy pomiaru konkretnego zjawiska, np. zakażeń szpitalnych, PRZY uwzględnieniu wyżej wyszczególnionych kryteriów, które mogą wpływać na wartość współczynnika zakażeń

Istnieje wiele technik dla uwzględnienia takiej korekty, ale żadna w pełni uwzględnia wszystkie kryteria dla danego pacjenta. Mamy jednakże określone techniki podstawowe (epidemiologiczne i kliniczne), dzięki którym można uzyskać rzetelne porównania wartości wskaźników.

D. Wiarygodność pomiaru

Jak wspomniano powyżej, wiarygodność związana jest z wielkością błędu, jakim obarczony jest pomiar przy pomocy wskaźników. Na przykład, pomiar z zastosowaniem „perfekcyjnie” wiarygodnego wskaźnika może być obarczony błędem rzędu 10 lub 0%. Większość wskaźników w ochronie zdrowia bywa obarczona błędem pomiaru przewyższającym 20%.

Skąd ten błąd? Oczywiście, sam wskaźnik może być skonstruowany w taki sposób, że „prosi się” o błąd. Definicje licznika i mianownika, a szczególnie niedostatecznie jasne określenie przypadków włączanych i wykluczanych, może prowadzić do problemów z wiarygodnością. W naszej dyskusji, dwa typy błędów są szczególnie istotne: błędy *systematyczne* i *losowe*.

Błędy systematyczne. Są to błędy uporczywie popełniane w trakcie procesu zbierania i przekazywania danych. Na przykład, do badania włączane są wszystkie przypadki readmisji, kiedy wymóg stanowi, iż ponownie hospitalizowani pacjenci mają zostać wyłączeni z mianownika. Czym skutkuje popełnienie takiego błędu?

Szpital, który popełnia błąd systematyczny może nadal porównywać swoje dane z innymi własnymi danymi zarejestrowanymi w innym okresie czasu. Dlaczego? Porównanie jest możliwe, gdyż względna różnica w wartości wskaźników, uwzględniająca systematyczny rodzaj błędu, może być nadal odnoszona do wyniku jaki uzyskuje ten szpital.

Jednak dany szpital nie będzie mógł porównywać wskaźników z innym szpitalem (ani ze średnią, jaką na przykład uzyskują szpitale uczestniczące w określonym badaniu czy projekcie) ponieważ inni najprawdopodobniej nie popełniają błędu systematycznego dla tych samych wskaźników.

Zatem, jeżeli popełniacie błąd systematyczny i wiecie na czym on polega (np. niewłaściwe wyłączenia dla mianownika) możecie NADAL wykorzystać wskaźniki do wewnętrznych porównań.

Błędy losowe. Te są bardziej kłopotliwe, gdyż trudno je przewidzieć - są nieprzewidywalne. Błąd losowy może dotyczyć jednego wskaźnika w danym miesiącu, ale może też już się nigdy nie powtórzyć. Ponadto, błędy mogą pojawiać się losowo w odniesieniu do różnych wskaźników. A zatem błąd losowy to taki, którego wystąpienia można oczekiwać, ale często nie można mu zapobiegać, gdyż trudno przewidzieć kiedy i jak do niego dojdzie. (gdyby był przewidywalny, byłby błędem systematycznym).

Jest i dobra wiadomość, a mianowicie że w inicjatywach takich jak PATH, znaczna liczba uczestników i pomiarów obarczonych błędami losowymi, ma NIEWIELKI wpływ na wartości wskaźników. Ponadto, obliczanie średniej (lub mediany) jeszcze bardziej minimalizuje wpływ błędów losowych (efekt wyrównywania). Tak więc, chociaż wiemy, że błędy losowe będą się pojawiać w badaniach, które wykorzystują wskaźniki dla porównań, błędy te NIE POWINNY skłaniać do zaniechania analizy porównawczej.

Rola wartości (value) w interpretacji wskaźników

Chociaż prawdziwa nauka jest obiektywna w swoich zastosowaniach, wskaźniki w ochronie zdrowia można określić jako „paranaukowe”. Przyczyną jest całe mnóstwo czynników, które ingerują w interpretację wskaźników, począwszy od niedookreślonej relacji pomiędzy procesem a wynikiem, a kończąc na niemożności zastosowania określonej metody dla uwzględnienia wszystkich aspektów związanych z pacjentem i zarządzaniem opieką.

Nie oznacza to jednak, że całość interpretacji musi być zgodna z oczekiwaniami lub wartościami dla danego szpitala, województwa czy kraju. W rzeczywistości, wiele wskaźników wykazuje pewien powszechnie akceptowany „kierunek” poziomu wartości np. wysoka wartość wskaźnika oznacza problem. Jest tak dlatego, że większość wskaźników skonstruowanych jest w sposób zmierzający do określenia odchylenia od wartości docelowej lub zmienności wokół niej. Pamiętajmy, że dyskutowaliśmy o „normach” i „punktach odniesienia” oraz o punktach docelowych. Gdy więc istnieją dowody (a zatem i podstawy naukowe), że dany współczynnik winien mieć określoną wartość, nie należy mówić „ale my uważamy, że jesteśmy inni”. Tak wygląda podejście oparte na dowodach naukowych, które prowadzą do określenia wytycznych postępowania klinicznego. A wytyczne nie powinny być modyfikowane w taki sposób, że istota zalecenia traci skuteczność.

Jednak przy braku wytycznych lub rzetelnych podstaw opartych na dowodach, interpretacja wskaźnika odbywa się na podstawie lokalnych oczekiwań i systemu wartości. Dobrym przykładem jest wskaźnik cięć cesarskich, ponieważ nie określono „idealnego” lub „normatywnego” poziomu dla tego wskaźnika, dotyczącego częstości wykonywania cięć pomimo jasnych wytycznych co do wykonywania cięć.

Waga i znaczenie wartości w interpretacji wskaźników zastaną omówione gdy będzie mowa o pojęciu ewaluacji, którą można zdefiniować jako „przypisywanie konkretnej obserwacji danej wartości”

EWALUACJA – Sztuka harmonizacji “wartości” i “oczekiwań”

Zacznijmy od ustalenia czym jest „ewaluacja”- jej definicji. Choć często jest określana jako „pomiar oddziaływania lub wpływania (na coś/kogoś)”, to w rzeczywistości jest „przypisywaniem wartości temu, co podlega obserwacji lub stanowi przedmiot badania”.

A zatem ewaluacja jest “przypisywaniem wartości temu, co badamy, w tym przypadku – poprzez dokonywanie pomiaru”.

Definicja ewaluacji jest bardzo istotna, gdyż wprowadza do rozważań pojęcia wartości i oczekiwań.

Implikacje dla “wartości” w trakcie ewaluacji

Wskaźniki prezentują określoną liczbę np. poziom 19% cięć cesarskich. Co decyduje o tym, czy uznać ten poziom za właściwy?

lub

Zgodnie z wytycznymi, wskaźnik dla antybiotykoterapii po zabiegach chirurgicznych wynosi 80%. Co to oznacza w świetle akceptacji przez chirurgów wprowadzonych wytycznych?

Te proste przykłady ilustrują wyzwanie, jakim jest uzyskanie zbiorowej zgody na to, jaki jest nasz cel, *PRZED* dokonaniem ewaluacji uzyskanych statystycznych pomiarów funkcjonowania. Wróćmy raz jeszcze do dwóch powyższych wskaźników:

a. Wskaźnik cięć cesarskich w szpitalu wynosi 19%. Jeżeli średnia dla tego wskaźnika w projekcie PATH wynosi 22%, czy oznacza to, że jest dobrze? I co znaczy „dobrze”? Czy na podstawie zgromadzonych danych dowiedzie się:

- czy wskaźnik cięć cesarskich był inny dla grupy kobiet w wieku 20-30 lat w porównaniu z grupą wiekową 30+? Jeśli tak, to czy mówi to cokolwiek o płodności, liczbie poprzednich porodów, schorzeniach współistniejących, powikłaniach lub preferencjach pacjentek dotyczących cięcia cesarskiego? Jeśli tak, to czy wychwycono którąś z powyższych zmiennych?
- jaki był stan zdrowia noworodków urodzonych przez cięcie cesarskie w porównaniu do dzieci urodzonych siłami natury? Na przykład, czy można powiązać liczbę porodów w szpitalu z liczbą przyjęć na oddział intensywnej terapii noworodka? Czy dane mogą wskazywać na to, że noworodki urodzone siłami natury są przyjmowane na intensywną terapię rzadziej, ale często są tam dłużej hospitalizowane? Czy można zebrać informacje o przyczynach pobytu na OIT noworodkowym i stosowanej tam terapii?
- czy koszt cięcia cesarskiego *PLUS* pobytu na oddziale intensywnej terapii noworodka był wyższy czy niższy niż koszt porodu naturalnego *PLUS* koszt pobytu noworodka na oddziale intensywnej terapii? Rozważając finansowe implikacje cięć cesarskich, być może okaże się, że cięcie kosztuje mniej niż poród naturalny, gdy uwzględnimy całkowity koszt opieki (nad matką i noworodkiem) podczas całego przebiegu porodu i opieki szpitalnej?
- o tym, że kobiety czytały w Internecie, że dzieci urodzone cięciem cesarskim, lepiej radzą sobie w szkole? O tym, że w społeczeństwie mogą występować bezpodstawne przekonania, na które system opieki zdrowotnej nie reaguje i nie prowadzi edukacji przyszłych matek w oparciu o wiedzę opartą na dowodach?

b. Wskaźnik(i) w szpitalu wykazuje(a), że stosowanie określonej antybiotykoterapii po zabiegu chirurgicznym wynosi 80%, podczas gdy średnia dla innych szpitali w projekcie PATH wykazuje poziom 90% dla tego miernika. Czy dane jakie zebrano pozwalają stwierdzić, czy:

- dane są zbierane w sposób zapewniający ich kompletność i wiarygodność? Proszę zauważyć, że kwestia kompletności danych w przypadku tego wskaźnika stanowi o wiele ważniejszy problem, niż w przypadku cięć cesarskich.

- chirurgom stale przypominano o znaczeniu tych wytycznych i konieczności ich stosowania, gdyż zostały one opracowane w oparciu o aktualną wiedzę medyczną i dowody naukowe (czyli są dokumentami *evidence-based*) ?
- istnieją zestawienia wyników leczenia, które umożliwiają porównanie rezultatów leczenia 80% pacjentów, u których zastosowano określone wytyczne dla antybiotykoterapii po zabiegu, w porównaniu z pozostałymi 20% ?
- w grupie 20% pacjentów pojawiły się powikłania, które generowały znaczne koszty opieki? Przykładowo, czy wystąpiły powikłania szpitalne wywołane poprzez “niewłaściwą antybiotykoterapię, których skutkiem był przedłużony pobyt w szpitalu?

Są to tylko przykładowe pytania, jakie należy sobie postawić, gdy mamy do czynienia ze wskaźnikami. W jaki sposób uwzględnić “wartości” w procesie ewaluacji?

Na podstawie obu powyższych przykładów widać, że jeden ze wskaźników (cięcia cesarskie) nie osiągnął idealnego czy też „prawidłowego” poziomu. W jego interpretacji należy uwzględnić wiele czynników tak klinicznych jak i socjologicznych. Ponadto, pewne wartości kultury systemu ochrony zdrowia mogą wskazywać, czy dana wartość wskaźnika na poziomie lokalnym, – nawet jeśli statystycznie odbiega od pozostałych wartości wskaźnika cięć cesarskich – może być zgodna z oczekiwaniami tak szpitala jak i hospitalizowanych pacjentek.

Drugi wskaźnik jest miernikiem niejako “pomostowym” ponad luką pomiędzy wytycznymi opartymi na dowodach a istniejącą praktyką lokalną. W tym przypadku, jeśli nie uda się nawiązać współpracy z zespołem chirurgicznym, wartość tego wskaźnika będzie bez znaczenia,. A zatem, nim skupimy się na analizie i interpretacji owego poziomu 80% przestrzegania wytycznych, należy się zastanowić dlaczego w pozostałych 20% przypadków nie postępowano zgodnie z wytycznymi. Wartość wskaźnika zmieni się tylko wtedy, jeżeli – w oparciu o zrozumienie korzyści stosowania wytycznych – zmieni się lokalna praktyka.

Na zakończenie proponuję, by spojrzeć na ewaluację jako proces obejmujący zarówno lokalną praktykę i zwyczaje, oczekiwania dotyczące “właściwego postępowania”, wytyczne postępowania klinicznego oparte na dowodach, pozwalające na porównania z innymi krajami, jak również proces, który podkreśla znaczenie zbierania danych na temat kluczowych zmiennych (klinicznych, demograficznych, socjologicznych), kształtujących każdą inicjatywę poprawy jakości opieki, a efekty tego ujrzycie w formie zmieniających się wartości wskaźników.

Akredytacja i rola ciągłego pomiaru przy pomocy wskaźników

Każdy usługodawca w sektorze usług społecznych odnosi korzyści, wynikające z uznania jego odpowiedzialności za świadczone usługi i dostarczane produkty. Powszechnie stosowaną metodą zapewnienia, że usługi świadczone są przez wykwalifikowane jednostki lub organizacje jest legitymizacja działalności w postaci licencji, której wyższą formą jest akredytacja.

Akredytacja oznacza ocenę dokonywaną przez organizację zewnętrzną, która obejmuje przegląd misji i strategii, celu i uzyskiwanych wyników działania. Ponieważ opieka nad pacjentem oraz promocja zdrowia są podstawowymi elementami opieki społecznej państwa i stanowią dobro społeczne, akredytacja winna być postrzegana jako systematyczny proces, który zapewnia ogółowi społeczeństwa usługi i produkty dobrej jakości.

Jednak akredytacja dla opieki zdrowotnej nie jest tak łatwa jak legitymizacja i uznanie w innych sektorach. Stanowi o tym szereg względów, z których dwa najważniejsze można określić następująco:

- a. Wiedza jest domeną świadczącego opiekę, ponieważ „klient lub użytkownik” często nie potrafi określić, czy dana usługa była prawidłowa, ani też, czy świadczone ją z najwyższą skutecznością. Nie tak, jak, przykładowo w kwestii mechanika samochodowego, którego kwalifikacje potrafimy ocenić, jak też to, czy prawidłowo zdiagnozował usterkę i ją naprawił.
- b. Brak jest dostatecznych danych na temat usług zdrowotnych, które byłyby zrozumiałe dla „klientów/użytkowników”.

Oba powyższe aspekty wskazują na konieczność zastosowania oceny zewnętrznej w systemie ochrony zdrowia.

Dlaczego pomiar funkcjonowania stanowi integralny element akredytacji?

Po prostu dlatego, że nie sposób poddać prawidłowej ocenie funkcjonowanie i działalność bez systematycznego, wiarygodnego i nieustannego zbierania danych. Sprostanie temu wyzwaniu nie jest jednak łatwe, jeżeli dla oceny stosuje się wyłącznie wskaźniki. Dlaczego?

Ponieważ akredytacja dotyczy co najmniej trzech wymiarów funkcjonowania organizacji jako funkcjonalnej całości: struktury, procesów świadczenia opieki oraz rezultatów dotyczących przede wszystkim zaspokajania zdrowotnych potrzeb pacjentów.

Wypowiedzi typu „nasi pacjenci są inni” i „mamy własne sposoby postępowania” to najczęściej stosowane uzasadnienia odnoszące się do obserwowanych różnicowań w zakresie struktury, procesu lub wyniku, odnotowywanych podczas oceny ilościowej lub przy zastosowaniu wskaźników. Rzeczywiście, nie sposób ocenić całości funkcjonowania szpitala przy pomocy samych wskaźników; nie ma też pojedynczego miernika, za pomocą którego można by ocenić całość działalności szpitala .

Podobnie nie można przyjąć jednostkowego trybu postępowania w odniesieniu do skarg, uwag i wniosków pacjentów.

Można natomiast zidentyfikować określoną liczbę najczęściej stosowanych procedur i usług zdrowotnych, występujących w większości szpitali, niezależnie od położenia geograficznego, czy lokalnej kultury świadczenia opieki. Ponadto, dotychczasowe doświadczenia wykazują, że skoncentrowanie uwagi na wybranych wskaźnikach służy nie tylko ilościowemu określeniu funkcjonowania szpitala, ale także okazuje się pomocne przy wprowadzaniu zmian dla poprawy w kulturze danej organizacji poprzez sam proces dokonywania pomiaru, oceny i porównań z określonymi wartościami lub wiarygodnymi punktami odniesienia. Stanowi to zasadniczy aspekt prowadzenia pomiaru funkcjonowania/działalności w ochronie zdrowia, który może być promowany przez system akredytacji, wypełniając społeczny mandat poprawy tej formy pomocy socjalnej, z której każdy z obywateli będzie kiedyś korzystał.

Względy strategiczne

Nie wykazano dotychczas, by akredytacja sama w sobie przyczyniała się bezpośrednio do poprawy wyników opieki. Piśmiennictwo i przeprowadzone badania wskazują, że uzyskanie akredytacji, choć zapewnia społeczne uznanie przyjęcia odpowiedzialności za świadczone usługi zdrowotne, to jednak nie daje absolutnej pewności, że akredytowana organizacja systematycznie monitoruje wyniki opieki, poddając je wewnętrznej analizie czy ocenie; nie gwarantuje także systematycznego zbierania danych dla poprawy i lepszej opieki.

W świetle powyższych rozważań, wydaje się, że w trakcie dyskusji na temat motywowania i gratyfikacji za uzyskiwane rezultaty opieki zdrowotnej, pojawia się unikatowa sposobność, by uznać akredytację za integralną część strategii rozliczania odpowiedzialności w sektorze ochrony zdrowia, taką, która realizuje społeczne oczekiwania, stanowiące nieodzowny element każdego systemu finansowania usług opartego na motywacji.