

Traduzione a cura di:
Infermiere DANIELA MOSCI e MARIA CRISTINA ROBB
in collaborazione con D.D.S.I. Paolo Chiari
Centri studi EBN - Direzione Servizio Infermieristico e Tecnico
Azienda Ospedaliera di Bologna – Policlinico S.Orsola-Malpighi
Via Massarenti, 9
40138 Bologna, Italia

Tel. e Fax. 051 6363049

E-mail: ebn@orsola-malpighi.med.unibo.it

Web: http://www.med.unibo.it/reparti_servizi/servinfer/homepage.html

I FONDAMENTI DELLA MISURAZIONE QUANTITATIVA

Lo scopo principale degli editoriali di EBN è di equipaggiare i lettori delle necessarie abilità per avvicinarsi in modo critico agli studi di ricerca primaria e fornire una descrizione più dettagliata di alcuni dei problemi metodologici che sorgono nei lavori che riassumiamo. Nel numero di luglio 1999 di EBN l'editoriale esplorava il concetto di campionamento. In questo numero forniremo un'introduzione di base alla misurazione quantitativa dei risultati sanitari, che possono valutare gli studi di trattamento, di eziologia, di prognosi, di diagnosi e di valutazione economica. Esempi di esiti sanitari sono la pressione arteriosa, la qualità della vita, la soddisfazione dei pazienti e i costi.

La salute può essere misurata in molte maniere; i vari aspetti della salute che possono essere misurati sono chiamate variabili. Per esempio nello studio sul trattamento di Dunn et al. in questo numero di EBN (p. 117), gli interventi (conosciuti come variabili indipendenti) sono l'attività fisica e il benessere cardiorespiratorio. In uno studio su un trattamento le variabili indipendenti sono quello che sono sotto il controllo degli investigatori e le variabili dipendenti sono gli esiti che possono essere influenzati dalle variabili indipendenti. In uno studio sull'eziologia gli investigatori fanno i conti con la variazione naturale tra entrambe le variabili e cercano una relazione tra le due variabili. Per esempio, quando vogliamo determinare se fumare provoca il cancro del polmone, fumare è la variabile indipendente e il cancro del polmone è la variabile dipendente. Nell'abstract incluso in EBN le variabili indipendenti sono identificate sotto la sezione "intervento" per gli studi sui trattamenti e sotto "fattori di valutazione del rischio" per gli studi sull'eziologia. Le variabili indipendenti sono identificate sotto la sezione "delle misure degli esiti principali".

TIPI DI VARIABILI

Le variabili possono essere classificate come nominali, ordinali, d'intervallo o di rapporto. Le variabili nominali (categoriche) sono semplicemente chiamate categorie. Alcune variabili nominali (di cui si parla come variabili dicotomiche) hanno solo due valori possibili, come il sesso (uomini o donne), la sopravvivenza (morto o vivo), o se una caratteristica è presente o assente (per esempio diabete: sì o no); altre variabili possono avere parecchi valori possibili come la razza (bianca, nera, ispanici e altri). Il numero reale delle categorie può essere determinato dal ricercatore; per esempio la razza può essere definita da due opzioni (nera o non nera) o da parecchie possibili opzioni. La gerarchia non è presumibile con i dati nominali - cioè, essere vivo non è due volte meglio che essere morto (nonostante la maggior parte dei pazienti potrebbe discutere con noi su questo argomento). Al contrario, le variabili ordinali sono serie di categorie "ordinate". Per esempio, ai pazienti è spesso richiesto di dare un punteggio alla severità del loro dolore su una scala da 0 a 10, dove 0 è assenza di dolore e 10 è un dolore insopportabile e atroce. Sebbene possiamo sapere con sicurezza che il dolore a cui è stato dato un punteggio di 8 è peggiore di quello valutato 5, non conosciamo veramente quanto questi due punteggi differiscano perché non conosciamo la dimensione dell'intervallo tra ogni punteggio. Le scale ordinali sono state usate anche per graduare la severità delle ulcere da decubito e per classificare lo stadio di varie tipologie di cancro (per esempio stadio I, II o III). Le variabili di intervallo consistono in una serie ordinata di categorie, con il requisito aggiuntivo che le categorie formino una serie di intervalli che sono tutti esattamente della stessa dimensione. Quindi, le differenze tra una temperatura di 37°C e di 38°C è 1°C e tra 38°C e 39°C è ancora 1°C e così via. Comunque una scala di intervallo non ha un punto zero assoluto, che

indica l'assenza completa della variabile in esame, a causa del fatto che non c'è il punto zero assoluto su una scala di intervallo, i rapporti di valore non sono significativi - cioè due valori non possono essere comparati con la pretesa che uno sia due volte più grande dell'altro. Una variabile di rapporto ha tutte le caratteristiche di una variabile di intervallo, ma aggiunge il punto zero assoluto, che indica assenza (completa) della variabile misurata. Il vantaggio di un punto zero è che il rapporto tra i numeri nella scala riflette il rapporto di grandezza per la variabile in esame: 100°C non è due volte più caldo di 50°C (intervallo dei dati), ma 100 cm. è due volte più lungo di 50 cm. e una pulsazione di 80 battiti al minuto è due volte le pulsazioni di 40 battiti al minuto (rapporto dei dati).

La ragione per accentuare le differenze tra i tipi di variabili è che il tipo di variabili impone, in larga misura, il metodo dell'analisi statistica usata dai ricercatori. Non ha senso discutere di media o valore medio di dati nominali o ordinali perché sono categorie. Cioè, la nozione di media di sesso o di razza non ha senso; il conteggio di frequenza del numero di individui in ogni categoria è comunque utile. Diversamente, la pressione arteriosa media (variabile di rapporto) in un campione di pazienti è più significativa che un conteggio del numero di pazienti con ciascuna misurazione della pressione arteriosa. In un successivo editoriale EBN esploreremo i modi differenti di descrivere e analizzare i dati.

PROBLEMI NELLA MISURAZIONE

È importante ricordare che la maggior parte delle misurazioni nella ricerca sanitaria incorporano parecchi elementi: il valore "reale" o vero della variabile che è stata misurata; la variabilità della misura; l'accuratezza dello strumento con il quale stiamo misurando e forse la posizione del paziente o le abilità e le aspettative della persona che sta' compiendo la misurazione. Alcuni di questi elementi sono entro il controllo del misuratore (per esempio assicurarsi che una pesa sia a "0" prima di pesare qualcuno), mentre altri elementi non lo sono (per esempio la pressione arteriosa di un paziente varia nella giornata; quindi i ricercatori provano a valutare la pressione arteriosa alla stessa ora ogni giorno).

Alcune misure sono più oggettive che altre e sono meno probabilmente influenzate dall'errore umano o dai bias. Esempi di misure oggettive includono tutte le cause di mortalità (cioè se un soggetto è "vivo" o "morto") e la concentrazione sierica del colesterolo. Diversamente, le misure soggettive possono essere influenzate dalla percezione della persona che sta' facendo la misurazione (per esempio l'autovalutazione del dolore da parte dei pazienti). La maggior parte dei questionari fatti con carta e penna sono misure soggettive. Il Beck Depression

Inventory for Primary Care descrive, nello studio sulla diagnosi di Steer e al. presente in questo numero (p. 126), un esempio di un questionario soggettivo fatto con carta e penna. Conteggi frequenti come la prevalenza o l'incidenza, sono spesso usate quando vogliamo conoscere l'estensione di una malattia o di una condizione in una popolazione. Altri possono essere interessati agli effetti benefici o dannosi di un intervento, come la differenza in percentuale delle malattie a trasmissione sessuale dopo un intervento comportamentale effettuato su donne minorenni (il trattamento studiato da Shain e al. p. 121).

QUALI PROBLEMI DI MISURAZIONE DOVREI CERCARE QUANDO LEGGO UN ARTICOLO?

LE MISURE SONO VALIDE E AFFIDABILI?

Queste sono due proprietà importanti dal punto di vista critico.

L'affidabilità riguarda il grado con cui una misura dà lo stesso risultato due volte (o di più) nelle stesse circostanze e che questa possa essere messa in relazione con la misura usata o le persone che la usano. Per esempio, se la pressione arteriosa di un paziente è misurata ogni 4 minuti sullo stesso braccio, dalla stessa infermiera, e il paziente non è sottoposto a nessun intervento come una attività o una misurazione, dovrete aspettarvi di avere le stesse letture dello sfigmomanometro. L'estensione con la quale letture ripetute sono simili viene chiamata affidabilità. La valutazione delle similarità di letture ripetute prese dalla stessa infermiera fornisce una misura di affidabilità interna. Dovreste anche sperare che due infermiere diverse, che eseguono la misurazione della stessa pressione arteriosa del paziente e nelle stesse circostanze, dovrebbero fornire letture simili: l'estensione con la quale la lettura delle due diverse infermiere sono simili è conosciuta come affidabilità interna.

La validità è l'abilità dello strumento di misurazione di misurare accuratamente quello che si intende misurare. Ci sono molti differenti tipi di validità ma uno dei più importanti è la validità legata ai criteri che richiedono il confronto di una misura data con un gold standard o con la migliore misura esistente della variabile. Nello studio di Steer e al. (p. 126) i risultati ottenuti dal Beck Depression Inventory for Primary Care è stato comparato con i risultati di una intervista standardizzata basata sui criteri del DSM-IV e condotta da un medico. I risultati dell'intervista sono stati considerati il gold standard. Altri esempi di gold standard sono la lettura diretta della pressione venosa centrale per la misurazione della pressione arteriosa mediante uno sfigmomanometro; la concentrazione sierica degli ormoni per i risultati di un test delle urine di gravidanza.

LA MISURA È SOGGETTA A BIAS?

Ci sono diverse fonti potenziali di bias. Non è importante ricordare come si chiamano, ma dovreste essere in grado di riconoscere le fonti di bias in uno studio. Un modo in cui possono verificarsi dei bias in uno studio è quando il fornitore di assistenza sanitaria, i pazienti e i raccoglitori dei dati che partecipano ad uno studio su un intervento non sono “ciechi” o “mascherati” rispetto alla somministrazione del trattamento. In un mondo ideale, gli studi dovrebbero essere a triplo cieco - cioè il fornitore di assistenza sanitaria che somministra l'intervento, il paziente e lo staff di ricerca che misura gli esiti non dovrebbero conoscere quale trattamento il paziente ha ricevuto. Sebbene il triplo cieco sia possibile nelle sperimentazioni randomizzate che valutano un nuovo farmaco è veramente molto difficile applicarlo nella valutazione della maggior parte degli interventi infermieristici. Spesso, né l'infermiera che somministra l'intervento né il paziente che lo riceve possono essere all'oscuro (per esempio le infermiere conoscono che stanno fornendo un intervento educativo al paziente e i pazienti sanno che lo stanno ricevendo). In tali studi è spesso possibile tenere all'oscuro la persona che misura gli esiti. Assicurandosi che la persona che misura l'esito sia tenuta all'oscuro dell'allocatione dei pazienti nei gruppi, i ricercatori tentano di minimizzare il bias che potrebbe essere introdotto da un aggiustamento inconsueto fatto dai valutatori, se essi sono consapevoli dell'allocatione dei pazienti nei gruppi. Per esempio, nello studio di Dunn e al. (p 117), che confronta due interventi per aumentare l'attività fisica, le persone che valutano la pressione arteriosa, la frequenza cardiaca e il grasso corporeo non conoscono quale intervento abbiano effettuato i partecipanti. Se essi lo avessero conosciuto, questo potrebbe avere influenzato la loro percezione quando dovevano effettuare la misurazione, in particolare se hanno un'opinione precisa circa quale sia l'intervento più efficace. Allo stesso modo, i partecipanti che riportano i loro livelli di attività potrebbero alterare la loro relazione comportamentale a seconda che si divertano o che sperino di venire collocati in un altro gruppo. Iniziando con questo numero di EBN specificheremo nella descrizione del disegno se lo studio non sia cieco o sia a singolo o doppio cieco e chi è cieco.

Un altro tipo comune di bias è il bias della desiderabilità sociale, nel quale le risposte delle persone alle domande possono riflettere i loro desideri di riportare in misura minore le loro abitudini socialmente sfavorevoli, come il numero di sigarette fumate, l'uso di droga illegale o di pratiche sessuali non protette. Diversamente, le persone

possono sovrastimare quello che percepiscono essere una pratica socialmente desiderabile, come la partecipazione agli esercizi fisici o l'assunzione giornaliera di frutta e vegetali.

Un terzo tipo di bias è il bias di richiamo, il quale dimostra che la memoria umana è fallibile, come evidenziato dalle relazioni sull'uso della cintura di sicurezza 5 anni fa o dall'assunzione di fibra il mese scorso che non sono così accurate come le misurazioni correnti o prospettive, dove l'uso delle cinture di sicurezza o il diario dietetico sono registrate in modo giornaliero.

Gli investigatori spesso usano strategie per tentare di superare questi bias potenziali. Queste strategie includono il disporre di valutatori degli esiti che non conoscono lo scopo di uno studio né quale intervento riceve il paziente: facendo completare i questionari di auto-valutazione ai partecipanti ad uno studio in sede privata; assicurando che le loro risposte a domande potenzialmente imbarazzanti e sensibili siano confidenziali; raccogliendo le informazioni su una base prospettiva (cioè come è successo), piuttosto che su una base retrospettiva (storicamente).

CONCLUSIONI

In un riassunto, i lettori di una relazione di ricerca hanno necessità di considerare il tipo di misure che sono state usate, l'affidabilità e la validità delle misure e i metodi usati per minimizzare i bias nella misurazione degli esiti. Questi sono alcuni elementi da considerare quando selezioniamo gli studi per la stesura degli abstract su EBN. Nel prossimo numero del giornale l'editoriale di EBN sarà indirizzato a come gli esiti degli studi sono analizzati e l'appropriatezza dei test statistici per il tipo di dati raccolti.

Bibliografia: vedi articolo originale

DONNA CILISKA, RN, PHD
and ALBA DiCENSO RN PhD
School of Nursing,
Faculty of Health Sciences
McMaster University,
Hamilton, Ontario, Canada

NICHY CULLUM RN PhD
Centre for Evidence Based Nursing, Department of
Health Studies,
University of York,
York, UK

