

## L'infermiere e la statistica: il metodo statistico e il ciclo Problema-Piano-Dati-Analisi-Conclusioni (PPDAC)

Carlo Di Pietrantonj<sup>1</sup>, Ines Basso<sup>2</sup>

<sup>1</sup>Servizio di Riferimento Regionale per l'Epidemiologia - SEREMI- ASL Alessandria. Italia

<sup>2</sup>Centro PhD Student Infrastruttura, Ricerca, Formazione, Innovazione. Azienda Ospedaliera SS. Antonio e Biagio e Cesare Arrigo, Alessandria, Italia. Dipartimento di Medicina Traslazionale. Università del Piemonte Orientale. Novara Italia

Corrispondenza: karldp@alice.it

### RIASSUNTO

La crescente necessità di un continuo aggiornamento professionale e l'estendersi del ruolo degli infermieri in nuovi ambiti richiede una sempre più approfondita competenza statistica per comprendere e valutare la ricerca scientifica sanitaria e saper pianificare e condurre ricerche. Le competenze riguardo i metodi statistici e le tecniche di organizzazione dei dati, per quanto possano apparire lontane dalla professione dell'infermiere sono in realtà costitutive del capitale intellettuale infermieristico. In questo breve lavoro descriveremo l'inquadramento logico del processo inferenziale come processo di tipo induttivo, che conduce a conclusioni probabilistiche il cui grado di affidabilità sarà determinato dalle "garanzie" che le premesse potranno fornire alle conclusioni, con il limite di essere sempre vulnerabili a nuove evidenze. Risultati poco vulnerabili a nuove evidenze, ovvero ricerche statistiche di buona qualità sulle cui stime poter orientare l'organizzazione sanitaria o poter praticare scelte terapeutiche importanti per i pazienti, possono essere ottenute solo attraverso una accurata pianificazione. Illustreremo il ciclo PPDAC come strumento per la pianificazione metodica della ricerca statistica.

### The nurse and statistics: the statistical method and the cycle Problem-Plan-Data-Analysis-Conclusions (PPDAC)

#### ABSTRACT

The growing need for continuous professional updating and the extension of the role of nurses into new areas requires an in deep statistical competence to understand and assess scientific health research and to be able to plan and conduct new research. The skills regarding statistical methods and data organization techniques, however distant they may appear from the nursing profession, are actually constitutive of the nursing intellectual capital. In this short paper we will describe the logical framing of the inferential process as an inductive process, which leads to probabilistic conclusions whose degree of reliability will be determined by the "guarantees" that the premises can provide to the conclusions, with the limit of always being vulnerable to new evidence. Results that are not vulnerable to new evidence, that is, good quality statistical research on whose estimates it is possible to guide the health organization or to be able to make important therapeutic choices for patients, can only be obtained through careful planning. We will illustrate the PPDAC cycle as a tool for methodical planning of statistical research.

#### INTRODUZIONE

Negli anni si è assistito da parte della professione infermieristica sia alla crescente

necessità di un continuo aggiornamento professionale, che richiede una sempre più approfondita competenza statistica per valutare la ricerca sanitaria pubblicata sulle riviste

scientifiche, sia all'estendersi del ruolo degli infermieri in ambiti professionali nei quali la preparazione in statistica e nella gestione dei dati risultano essere indispensabili. Queste competenze, che vanno dall'organizzazione metodica dei dati e delle informazioni, all'uso coscienzioso dei metodi e delle tecniche statistiche, sono costitutive del capitale intellettuale infermieristico come risulta evidente dalla storia professionale della fondatrice dell'infermieristica moderna (AA.VV 202; Betts e Wright, 2020; Hassmiller, 2021; Kopf, 1916; Nuttall, 1983).

Attraverso questo contributo vorremmo illustrare i concetti cardine della gestione informatica e dell'elaborazione statistica dei dati di interesse infermieristico, senza la pretesa di insegnare a fare lo statistico o l'informatico né di risolvere particolari problemi. Invece, vorremmo sottolineare ciò che il fruitore di ricerche condotte da altri, oppure chi aspira a impadronirsi dei metodi statistici devono aver presente.

**Scopi dell'indagine statistica in sanità:** in sintesi si possono riconoscere due obiettivi generali della ricerca statistica, il primo è quello di trarre le migliori informazioni possibili dai dati disponibili, ovvero ottenere informazioni applicabili al di là della limitata esperienza empirica che le ha prodotte; il secondo è quello di permetterci di prendere decisioni, ovvero consentire scelte razionali in condizioni di "incertezza". Il procedimento statistico che ci permette di perseguire questi due obiettivi generali e che ci consentirà di dare sostanza al termine "incertezza", viene generalmente denominato "inferenza statistica".

**Inferenza, deduzione, induzione:** l'inferenza statistica è quel procedimento che a partire da un insieme limitato di osservazioni (campione) permette di trarre conclusioni che possono essere generalizzate, ovvero "applicabili più in generale", a tutti quei soggetti (popolazione) che presentano le medesime caratteristiche del limitato gruppo realmente osservato. Il termine deriva da latino *in-fero* con il significato di portare dentro, arrecare, concludere, "trarre, partendo da una determinata premessa o dalla constatazione di un fatto, una conseguenza, un giudizio, una conclusione. Inferire una verità da un'altra" (Treccani 2022a). In termini filosofici l'inferenza è: "ogni forma di ragionamento con cui si dimostri il logico conseguire di una verità da un'altra" (Treccani 2022b).

Si riconoscono due forme principali di ragionamento, quello *deduttivo* definito come: "il processo logico nel quale, date certe premesse e certe regole che ne garantiscono la correttezza, una conclusione consegue come

*logicamente necessaria*" (Treccani 2022c), e quello *induttivo* ovvero: "il procedimento logico, opposto a quello della deduzione, per cui dall'osservazione di casi particolari si sale ad affermazioni generali" (Treccani 2022d). Il ragionamento deduttivo consente di formulare argomentazioni le cui conclusioni sono "sempre vere" ogni volta che sono soddisfatte le premesse. Il ragionamento deduttivo, poiché le conclusioni sono contenute nelle premesse, consente inferenze certe. Possiamo definire la "validità" di una argomentazione deduttiva come segue: "un argomento deduttivo è valido quando le sue premesse, se vere, forniscono ragioni decisive (conclusive) per la verità delle conclusioni" (Copi e Cohen, 1999). Per la logica deduttiva è quindi cruciale chiarire la relazione fra premesse e conclusioni in modo da distinguere gli argomenti validi da quelli "invalidi", quando dalla verità delle premesse non discende la verità della conclusione.

Il procedimento induttivo, invece, si distingue da quello deduttivo proprio per la differente natura della relazione fra premesse e conclusioni. L'argomentazione induttiva non richiede che le sue premesse forniscano ragioni decisive per la conclusione, ma soltanto che forniscano "qualche" sostegno per la conclusione (Copi e Cohen 1999). Pertanto gli argomenti induttivi non sono né validi né invalidi, tuttavia possono essere valutati come migliori o peggiori, forti o deboli, a seconda del grado di garanzia che le premesse danno alla conclusione (Copi e Cohen, 1999). Quindi il procedimento consente di ottenere conclusioni che contengono qualche cosa in più rispetto alle premesse, tuttavia, il metodo induttivo fornisce "inferenze incerte" o "probabili". Questa probabilità non solo è questione di grado, ma può dipendere da altre circostanze. Infatti l'aggiunta di nuove premesse (vere) può alterare la probabilità della conclusione rafforzandola o indebolendola, e nei casi estremi capovolgendone il senso. Quest'ultima è una seconda notevole distinzione fra argomentazioni deduttive e induttive; infatti, le argomentazioni deduttive non vengono perturbate da ulteriori premesse vere, in altri termini non sono vulnerabili di fronte a nuove evidenze.

L'inferenza statistica è di tipo induttivo, spesso indicata come "induzione statistica", poiché conduce a conclusioni (risultati), in linea di principio, applicabili al di fuori della limitata esperienza empirica su cui si basano; tuttavia, tali conclusioni non sono certe, sono probabilistiche il cui grado di attendibilità sarà determinato da un lato dalle "garanzie" che le premesse, le ipotesi di lavoro e i dati empirici potranno fornire alle conclusioni, dall'altro, dal grado di "accuratezza" e "precisione" con il quale tutte le premesse rilevanti/pertinenti sono state individuate e incluse prima di dare corso

al processo inferenziale. Quindi, le conclusioni portate dall'inferenza induttiva non solo possono essere alterate da nuove informazioni pertinenti non considerate (errore di selezione), ma anche dalla precisione con la quale le informazioni sono state rilevate (errore di misurazione), ovviamente nelle "nuove evidenze" che possono minacciare le conclusioni vi sono anche quelle premesse ritenute vere che in un secondo tempo risultano false.

Risultati attendibili, poco vulnerabili a nuove evidenze, ovvero: ricerche statistiche di buona qualità i cui risultati consentono di orientare l'organizzazione sanitaria per la tutela della salute dei cittadini; di poter praticare scelte terapeutiche importanti per i pazienti, possono essere ottenute solo attraverso una accurata pianificazione.

### LA PIANIFICAZIONE DI UNA RICERCA STATISTICA

La spinta a intraprendere una ricerca è alimentata dal desiderio di trovare una risposta a domande interessanti riguardo a ciò che osserviamo. Ad esempio, *"quanto sono efficaci gli interventi infermieristici per far cessare l'abitudine al fumo negli adulti?"* (Rice VH et al 2017), oppure *"qual è l'impatto degli infermieri specializzati in disabilità intellettiva nella assistenza alla persona con disabilità intellettiva?"* (Bur et al. 2021), e ancora *"quale è il grado di correlazione fra gli stili di leadership nei manager infermieristici e lo stress lavorativo degli infermieri e l'anticipo del turnover?"* (Pishgooie et al 2019).

Tutti questi "quesiti di ricerca", suscitati da problemi reali, richiedono al ricercatore di pianificare e condurre una ricerca che trovi una risposta attendibile (o se vogliono individui una soluzione), come descritto nel paragrafo precedente.

Le modalità di redazione di un protocollo di ricerca sono già state ampiamente descritte (O'Brien e Wright, 2022); (Bando e Sato 2015), tuttavia, non di rado, la parte della pianificazione statistica e la parte di gestione delle informazioni vengono organizzate successivamente e non come parte integrate del processo di pianificazione della ricerca. Questo avviene perché da un lato la statistica viene insegnata come una sequenza di strumenti matematici da applicare per risolvere di volta in volta particolari problemi, facendola apparire come un esercizio di applicazione di procedure specialistiche, dall'altro la gestione informatica dei dati viene percepita come estranea al metodo statistico e accessoria alla ricerca. L'indagine statistica non è data dall'applicazione di una serie di algoritmi e non

è separata dai metodi di gestione informatica dei dati, ed entrambe sono parte integrante del processo di ricerca e devono essere inquadrare in uno schema metodologico, che ne espliciti le fasi e ne consenta un'accurata pianificazione.

Una strategia di pianificazione della ricerca statistica che rappresenta il "metodo statistico" nella sua completezza è il ciclo *Problema-Piano-Dati-Analisi-Conclusioni* (PPDAC) (MacKay e Oldford, 2000); (de Smith, 2018); (Wild e Pfannkuch, 1999) costituito da cinque livelli nel quale ogni livello conduce al successivo ed è dipendente dal precedente, cioè ogni fase viene legittimata nel contesto delle fasi precedenti (MacKay e Oldford, 2000), ad esempio: senza aver definito con precisione i contorni del problema non ha significato pianificare uno studio, inoltre secondo (MacKay e Oldford, 2000) la complessità dell'analisi si riduce se "il problema" viene affrontato secondo tale struttura.

Il metodo PPDAC è sintetizzato in tabella 1; di seguito saranno esaminati più in dettaglio le diverse fasi.

La tabella 1 sembra suggerire un processo lineare nel quale in ogni fase discende dalla precedente, invece lo schema PPDAC è concepito in modo ciclico [figura 1], infatti la conclusione di uno studio può suggerire un nuovo problema da affrontare, ma soprattutto è concepito per essere utilizzato in modo iterativo fra fasi contigue, ovvero le scelte effettuate in una fase possono prevedere di rivedere o perfezionare quelle fatte nella fase precedente, prima di procedere alla fase successiva. Quindi l'utilità del PPDAC non è solo quello di fornire un completo modello pragmatico dell'indagine statistica, ma anche uno strumento di revisione critica degli stadi completati.

**Tabella 1.** - Fasi della pianificazione della ricerca; il ciclo PPDAC da (MacKay R J, Oldford R W, 2000, de Smith M J 2018) modificata.

Fase	Dettagli
<p><b>Fase1: Definizione del Problema o quesito di ricerca</b></p>	<p><b>Unità e Popolazione Obiettivo:</b> la popolazione (è il collettivo delle unità) verso la quale si vorrebbero “generalizzare” i risultati.  <b>Il quesito di ricerca (obiettivo dello studio):</b> ciò che vogliamo imparare dalla ricerca.  <b>Intento dello studio:</b></p> <ul style="list-style-type: none"> <li>• <b>Descrittivo- Esploratorio-</b> stimare attributi della popolazione (ad es prevalenza infezioni)</li> <li>• <b>Analitico-valutativo Predittivo:</b> accertamento di un nesso causale (ad es efficacia terapeutica; effetto dose-risposta).</li> </ul> <p><b>Variabili (esplicative e di risposta):</b> le variabili sono gli attributi delle unità statistiche.</p> <ul style="list-style-type: none"> <li>• <b>la variabile esplicativa (Determinante)</b> nel caso più semplice è dicotomica ovvero costituita da due categorie o livelli (ad es. esposizione/ non esposizione, oppure trattamento A / trattamento B)</li> <li>• <b>Confondenti, Modificatori d’effetto</b>, variabili esplicative il cui effetto sulla variabile di risposta si vuole controllare o eliminare</li> <li>• <b>la variabile di risposta (Outcome):</b> ad es l’esito di una terapia (dicotomica), il valore di un parametro ematochimico(continua).</li> <li>• <b>Ulteriori variabili esplicative</b> che possono aiutare a spiegare le variazioni della variabile di risposta al di fuori dell’intervento/esposizione.</li> </ul> <p><b>Attributi della popolazione:</b> caratteristiche di contesto e della popolazione (ad es il setting ospedaliero, epidemiologico, socio economico).  <b>Background</b> Che cosa già si sa del problema che si intende affrontare con la ricerca e che cosa potrebbe aggiungere questa ricerca</p>
<p><b>Fase 2: Piano Definizione del Disegno di studio</b></p>	<p><b>Definizione del modello in studio o disegno di studio:</b> la struttura logico formale delle relazioni fra variabili di risposta ed esplicative.  <b>Organizzazione delle variabili esplicative:</b> in termini di <i>determinanti, confondenti e modificatori di effetto</i>.  <b>Processi di misurazione e Parametri di Misura:</b> modalità di raccolta dati quantitativi. Metodi strumenti di misura. Valori di sintesi collettiva  <b>Misure d’effetto e di relazione:</b> differenze e rapporti di frequenze relative, differenze di medie, coefficienti di correlazione, OddsRatio  <b>Popolazione Osservabile:</b> l’insieme delle unità che è possibile raggiungere per determinare i valori delle variabili di interesse.  <b>Protocollo di campionamento:</b> tipo e schemi di campionamento  <b>Protocollo di raccolta dati:</b> la gestione dell’intero processo di raccolta e gestione dei dati, compresi gli aspetti tecnico informatici.</p>
<p><b>Fase 3: Dati-Struttura dei Dati</b></p>	<p><b>La struttura dei dati e tracciato record:</b> data base relazionali  <b>Vincoli di integrità dei dati per</b> garantire la coerenza interna delle informazioni  <b>Data Screening e Pulizia dei dati:</b> controlli per verificare la qualità e la completezza ed eventuali errori e loro rimozione</p>
<p><b>Fase4: Analisi Statistica</b></p>	<p><b>Analisi descrittiva:</b> sintesi numerica e grafica delle caratteristiche rilevanti del campione analizzato  <b>Analisi valutativa:</b> test d’ipotesi (significatività statistica), verifica del modello definito in fase di disegno di studio</p>
<p><b>Fase 5: Conclusioni</b></p>	<p><b>Discussione dei risultati:</b> Sintesi dei risultati rilevanti, sottolineare se essi sono conclusivi in accordo o in disaccordo con l’ipotesi iniziale, oppure se sono inconclusivi; discutere le similarità e differenze con risultati di altri studi simili; descrizione dei Potenziali limiti dello studio: indicare possibili bias che potrebbero limitare la validità dei risultati  <b>Conclusioni:</b> Sintesi delle conoscenze più importanti acquisite tramite lo studio; chiarire le implicazioni dei risultati per la pratica e per la ricerca futura.</p>

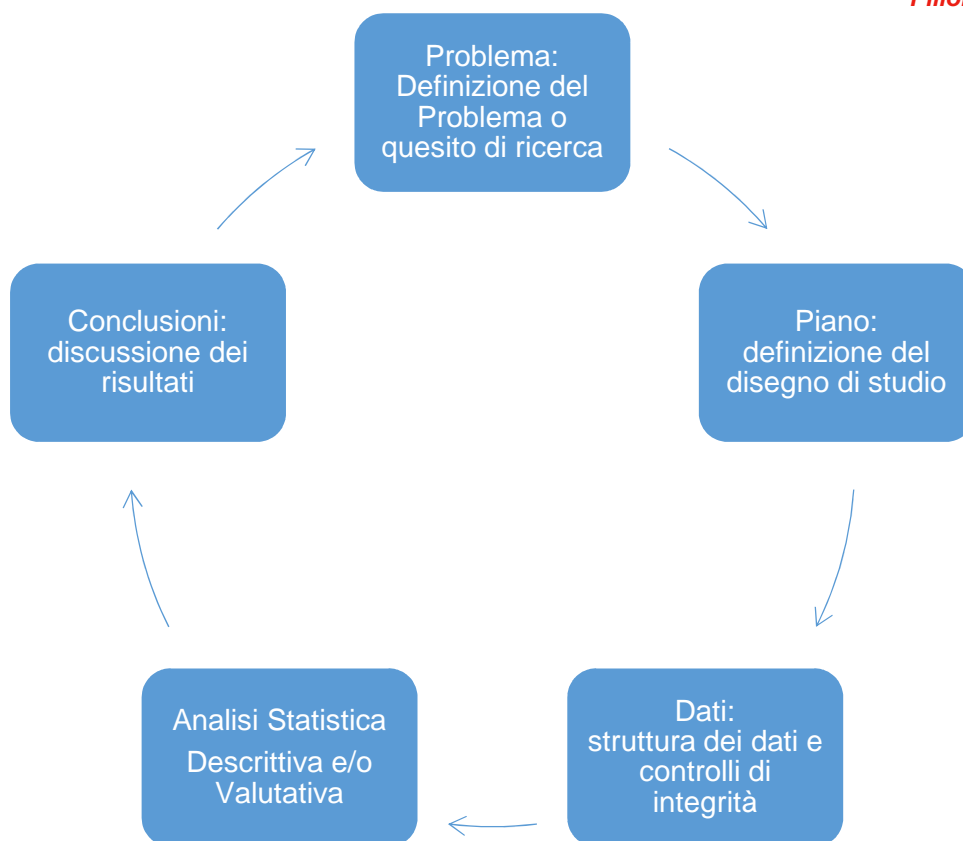


Figura 1. – Il ciclo PPDAC.

### FASE 1: Definizione del Problema o quesito di ricerca

Un buon quesito deve rendere immediatamente comprensibile il problema e che cosa si vuole imparare dalla ricerca (obiettivo della ricerca). Un quesito chiaro è un fattore decisivo per il successo della ricerca, sia in termini di raggiungimento degli obiettivi, sia in termini di impatto (de Smith, 2018), ovvero di comprensibilità e diffusione dei risultati fra gli *stakeholders*: operatori, amministratori, decisori, organizzazioni scientifiche e di cittadini. Una ricerca deve contribuire all'avanzamento delle conoscenze su di un problema specifico e a consentire scelte razionali. Individuare gli *stakeholders* e considerare il loro coinvolgimento attivo nella definizione del problema e nell'intero percorso di ricerca è strategico, poiché possono far emergere aspetti importanti del problema (de Smith, 2018); (Higgins *et al.* 2019). Una descrizione del problema e del contesto in modo chiaro e con terminologia appropriata aiuterà a guidare la progettazione in tutte le fasi successive. (MacKay e Oldford, 2000).

**Popolazione Obiettivo:** è la popolazione verso la quale la ricerca vorrebbe generalizzare i suoi

risultati; essa potrà essere definita sia in termini di collettivo di unità di osservazione che rispondono a certe caratteristiche (attributi) e/o in termini del processo (algoritmo) che ne definisce i criteri di appartenenza; ad esempio nel caso del quesito sull'efficacia dell'infermiere specialista per l'assistenza alle persone adulte con disabilità intellettiva, la popolazione obiettivo sono tutte le persone adulte con disabilità intellettiva. In questa fase è bene dare una definizione ampia a meno che non vi siano già ragioni per definire segmenti più specifici della popolazione obiettivo, come la presenza di altre condizioni di salute o patologie.

**Intento dello studio:** l'intento può essere descrittivo-esploratorio, analitico-valutativo o predittivo.

L'intento *descrittivo-esploratorio* è quello di descrivere fenomeni o eventi nella popolazione (ad es la frequenza di una malattia) e individuare quali strati o fattori sono potenzialmente collegati a sue variazioni. In altri termini lo scopo è quello sia di stimare la frequenza dei fenomeni e descrivere le loro caratteristiche, sia individuare quali sono i possibili fattori che potrebbero influenzarne la comparsa. Questo tipo di indagine risponde alle



domande: Chi? Che Cosa? Dove? Quando?

L'intento *analitico-valutativo* è quello di valutare l'esistenza di nessi causali, ad esempio: se l'esposizione a fattori specifici è associata (o non associata) allo sviluppo di specifiche malattie; stabilire quali interventi sono associati (o meno) a una migliore prognosi; se l'introduzione di una novità organizzativa in un servizio ha avuto l'effetto (o non ha avuto l'effetto) nel migliorare la soddisfazione e/o lo stato di salute degli utenti. In altri termini lo scopo è quello di raccogliere i dati in modo da poter decidere fra due possibili ipotesi. Infine, *l'intento predittivo*, di fatto una estensione dell'intento analitico-valutativo, è quello di stabilire quali valori assume la variabile di risposta, al variare dei valori assunti dalle variabili esplicative. Questi tipi di indagine rispondono alle domande: Come? Perché?

Esiste una sostanziale differenza fra l'intento esploratorio e l'intento analitico. Infatti, l'analisi esploratoria ha lo scopo di generare ipotesi, che dovranno essere vagliate con gli studi analitici, per questo motivo le analisi condotte con questo secondo intento sono anche dette *confermative* (sebbene il termine *confermativo* sembri sott'intendere una "dimostrazione di verità" ricordiamo che la natura induttiva dell'inferenza statistica fa sì che anche la più solida analisi può solo concederci risultati nel "crepuscolo della probabilità").

**Obiettivo dello studio:** deve far comprendere cosa si vuole imparare dalla ricerca. Dagli esempi di quesito citati emerge una sintassi che organizza i termini in modo da costruire una domanda che faccia cogliere immediatamente gli scopi della ricerca. Sugeriamo di utilizzare i criteri proposti dalla Cochrane (Cochrane Work 2022):

- [Intervento A] per intervenire (o prevenire) [il problema di salute] nella [popolazione/contesto].
- [Intervento A] confrontato con [intervento B] per ridurre [il problema di salute] nella [popolazione].
- Valutare se [Intervento/esposizione] riduce (aumenta) il rischio di [problema di salute] nella [popolazione esposta].
- Misurare la diffusione del [problema di salute] nella [popolazione/contesto].

Dopo l'obiettivo generale è possibile inserire gli obiettivi specifici e secondari; anch'essi dovrebbero essere elencati in modo conciso ed è buona norma non includerne troppi, poiché spesso portano a risultati imprecisi e poco definiti (Al-Jundi e Sakka, 2016).

Il criterio più noto per ricordare gli elementi di un quesito è incapsulato nell'acronimo PICO (Lazzari e Salvini et al 2015); (Higgins et al. 2019). Popolazione: *pazienti affetti da disabilità*

*intellettiva*; Intervento o esposizione al trattamento: *l'assistenza da parte di infermieri specializzati in disabilità intellettiva*; Comparatore o categoria di riferimento: *pazienti della medesima coorte con medesime caratteristiche degli "esposti/trattati", ma non assistiti da infermieri specializzati*; Outcome: l'esito, o variabile di risposta (*utilizzo dei servizi sanitari*).

Il quesito generale così formulato può anche essere usato anche come titolo del protocollo (Cochrane Work 2022) e del resoconto finale della ricerca, ciò aiuterà ad attirare i potenziali lettori (O'Brien e Wright, 2022). Inoltre, resoconti di ricerca e protocolli spesso vengono memorizzati in archivi informatici, pertanto un titolo strutturato come l'obiettivo facilita il recupero tramite i motori di ricerca.

Infine, formulato l'obiettivo è bene valutare quanto esso possa essere solido. A tal fine sono stati proposti cinque criteri, sintetizzati dall'acronimo FINER (Higgins et al. 2019), ovvero: *Fattibile, Interessante, Nuovo, Etico e Rilevante*. *Fattibile*: la domanda deve poter essere affrontata con le tecniche e le tecnologie disponibili, inoltre deve essere possibile reperire i dati e le informazioni pertinenti; *Interessante*: ogni ricerca richiede un impegno, pertanto deve essere interessante per chi la conduce; *Nuovo*: il lavoro deve colmare una reale lacuna nella conoscenza, il ricercatore deve essere consapevole di eventuali lavori simili attraverso una ricerca della letteratura rilevante, prima di iniziare il lavoro; *Etico*: se coinvolge direttamente i pazienti deve rispettare la privacy e i principi etici e non deve produrre risultati che amplino le disuguaglianze. Infatti, le domande di ricerca non sono neutre in termini di valore e il modo in cui viene affrontato un determinato problema può avere implicazioni politiche che possono comportare l'ampliamento delle disuguaglianze sanitarie (intenzionali o meno); *Rilevante*: i risultati devono essere importanti per gli stakeholders (Higgins et al. 2019).

**Le variabili di risposta ed esplicative** sono gli elementi del quesito di ricerca. In questa fase è necessario descriverle ed enunciare quale significato hanno considerando l'intento dello studio. Negli studi analitico-valutativi, volti a rilevare un nesso causale, è necessario distinguere fra variabile esplicativa e variabile di risposta. L'obiettivo di tali disegni è determinare l'effetto delle variazioni delle variabili esplicative (dette anche variabili indipendenti) sulle variazioni delle variabili di risposta (dette anche variabili dipendenti), ovvero se esiste un effetto delle prime sulle seconde. Mentre, quando l'intento è descrittivo-esploratorio vanno descritte quali sono le variabili di interesse che caratterizzano la popolazione al fine di

esplorare quali connessioni reciproche possono avere, senza dare una distinzione fra variabili esplicative e di risposta. La definizione più tecnico operativa delle variabili verrà data nella fase 2.

**Le variabili di risposta (Outcome)** esprimono l'evento o lo stato del quale si vorrebbe studiare l'occorrenza (la tendenza a presentarsi) nella popolazione, possono essere descritte usando variabili dicotomiche ovvero con due sole categorie, ad es infetto/non infetto, o variabili politomiche (molte categorie), ad es. stadio di una malattia, oppure variabili continue (misure) ad es. temperatura o pressione diastolica, o semi continue (variabili ordinali, discrete, conteggi o variabili a scala) come il numero dei globuli bianchi o gli indici di auto sufficienza (tabella 2). Nel caso si studi analitici su di un intervento (ad es nuovo farmaco o procedura vs standard) è importante descrivere oltre agli esiti attesi principali anche i possibili esiti secondari "desiderabili" o "indesiderabili" (ad es effetti collaterali) che si possono presentare associati all'esito, in genere già noti da studi precedenti.

**Le variabili esplicative** sono le variabili (fenomeni) che si presume esercitino un effetto sulle variabili di risposta. Alcune variabili esplicative sono di primario interesse e vengono dette determinanti, poiché rappresentano quei fenomeni per i quali si vuole indagare quanto le loro variazioni abbiano un impatto sulla variabile di risposta (outcome). Nel caso più semplice il determinante è dicotomico (es. intervento/non-intervento), quindi segmenta la popolazione obiettivo in due sottopopolazioni (livelli). L'obiettivo analitico-valutativo è quello di stabilire se la transizione del determinante dallo stato di "non-intervento" (livello di riferimento) a quello di intervento è collegata o meno ad una variazione dei valori (o categorie) della variabile di risposta (outcome). Proseguendo l'esempio sull'impatto degli infermieri specialisti la variabile esplicativa principale (il determinante) è dicotomica ovvero con due livelli: l'assistenza infermieristica specialistica e assistenza standard, mentre la variabile di risposta (outcome) è una variabile discreta, come ad esempio il numero di accessi ai servizi sanitari.

**Modificatori d'effetto e confondenti:** oltre a elencare e descrivere le potenziali variabili esplicative principali, vanno elencate le variabili che potrebbero interferire nella relazione determinante-outcome (il modificatore d'effetto), oppure essere responsabili di una apparente relazione determinante-outcome (il confondente). Nel nostro esempio la relazione determinante-outcome è la relazione fra tipo di assistenza e numero di accessi ai servizi sanitari.

La relazione determinante-outcome può non essere costante, ma dipendere dai livelli di un'altra variabile, ad es l'intensità della risposta ad un trattamento può dipendere dal livello di aderenza dei pazienti (compliance), oppure dall'età dei pazienti. I modificatori della relazione (o d'effetto) sono essi stessi determinanti (Rothman, 2007). D'altra parte una variabile può solo essere "apparentemente" esplicativa della variabile di risposta senza in realtà esserlo. In generale il fenomeno è dovuto alla presenza di una terza variabile detta "confondente", collegata alla presunta variabile esplicativa (nella popolazione generale) e simultaneamente determinante della variabile di risposta nel livello della popolazione "non esposta" della presunta variabile esplicativa (Rothman, 2007).

I confondenti possono "confondere" la relazione (determinante-outcome) sia mostrando una relazione fittizia, sia mascherando una relazione reale. Ad es la relazione apparente fra numero di figlie e cancro al seno è confusa dall'età, infatti l'età e il numero di figli sono associate nella popolazione generale. Inoltre l'età è determinante del cancro al seno anche nelle donne senza figli. In generale tutti i determinanti "conosciuti" per una malattia sono "potenziali" confondenti. Pertanto individuare questi fattori è importante per comprendere in dettaglio la natura della relazione fra variabili esplicative e variabili di risposta e soprattutto evitare di individuare alcuni fattori come esplicativi, quando in realtà non sono tali e viceversa. La conoscenza della letteratura rilevante sul problema consente di conoscere in anticipo i potenziali fattori confondenti e modificatori d'effetto.

Tipici fattori che possono giocare il ruolo di modificatore d'effetto o di confondente sono ad esempio: età, gravità della malattia, l'etnia, abitudini e stili di vita, modalità di erogazione dell'intervento, aderenza al trattamento, oppure alcuni **attributi della popolazione obiettivo** come: le condizioni di contesto: il setting ospedaliero o assistenziale, le modalità di organizzazione dell'assistenza sanitaria, le condizioni epidemiologiche o socio economiche, le caratteristiche del territorio (rurale, urbano, isolato).

**Background o base razionale della ricerca:** necessaria per presentare in un testo organico quanto illustrato sopra, ovvero cosa la ricerca aggiunge di nuovo rispetto alle conoscenze pregresse; fornire la rassegna di ciò che è noto del problema, riportando i risultati tratti dalla letteratura rilevante e fornire riferimenti ad altri studi simili congeniati per rispondere al medesimo quesito. Fornire una descrizione della condizione e del contesto, includendo informazioni, sull'epidemiologia, sulla clinica o

di sanità pubblica. Nel caso di uno studio su di un intervento deve descrivere come l'intervento potrebbe funzionare. Spiegare perché la domanda a cui si vuole dare risposta è importante. Nel caso di uno studio clinico fornire riferimenti sulla situazione corrente o pratica clinica. Descrivere come e da chi (*stakeholders*) i risultati potrebbero essere utilizzati.

## FASE 2: Piano della ricerca o disegno dello studio

Lo scopo di questa fase è definire il piano della ricerca o disegno di studio. Questa fase sarà più semplice se la fase precedente è stata svolta con cura. Tuttavia potrebbero essere necessari diversi passaggi fra fase 2 e fase 1 per definire sia un buon piano di ricerca, sia un chiaro quesito di ricerca.

**Disegno di studio:** è un modello della realtà che si intende studiare, ovvero una struttura logico formale nella quale viene chiarita la relazione fra la variabile di risposta e le variabili esplicative e il ruolo di quest'ultime all'interno del modello in termini di determinanti, modificatori d'effetto e confondenti. Inoltre, deve essere chiarita la popolazione che genera le osservazioni incluse nello studio, ovvero i criteri di eleggibilità e le modalità di selezione dei soggetti (tecniche di campionamento), il modo con cui si intendono misurare le variabili e la durata dell'osservazione. Il tipo di disegno di studio dipende se l'intento è esplorativo-descrittivo oppure analitico-valutativo. In sintesi esprime le strategie adottate dai ricercatori per rispondere al quesito di ricerca (Polit e Beck, 2014).

L'obiettivo di tutti i disegni di studio analitico-valutativi è quella di garantire le tre comparabilità: *la comparabilità delle popolazioni* individuate dai livelli della variabile esplicativa principale, ad es. sottopopolazione esposta vs non-esposta; *la comparabilità degli effetti*, tutti gli effetti estranei che possono agire sulle variazioni delle variabili di risposta, fra esposti e non-esposti e *la comparabilità delle misure*, l'uniforme accuratezza di rilevazione delle informazioni fra esposti e non-esposti.

Il disegno analitico-valutativo che garantisce al meglio tali comparabilità è quello sperimentale, dove il ricercatore decide l'esposizione e controlla le condizioni in cui l'esperimento è svolto. Mentre nei disegni analitici dove il ricercatore non decide l'esposizione, ma osserva gli eventi (studi osservazionali) è possibile controllare solo in parte le condizioni entro le quali tali eventi si svolgono. Quindi negli

studi analitici osservazionali è cruciale individuare tutte le potenziali variabili esplicative secondarie i cui effetti si vogliono "isolare" dall'effetto della variabile esplicativa di interesse. Gli studi osservazionali per quanto ben condotti, non forniscono assolute garanzie riguardo le tre comparabilità.

In ogni caso per tutti gli studi analitici l'affidabilità dei risultati sarà determinata dalla capacità della pianificazione del disegno di garantire le tre comparabilità e da quanto bene saranno evidenziate le possibili distorsioni, ovvero le differenze sistematiche fra le caratteristiche e gli attributi delle sottopopolazioni definite dai livelli della variabile esplicativa (determinante) che potrebbero influire sui risultati.

### Organizzazione delle variabili esplicative:

nei paragrafi precedenti abbiamo chiarito che vi è una importante distinzione fra le variabili esplicative: in termini di Determinanti, Modificatori d'Effetto e Confondenti. Nel caso degli studi descrittivi nei quali l'obiettivo è comprendere il grado relazione reciproca fra le variabili, non è necessaria una distinzione così fine. Mentre, negli studi analitici l'obiettivo è testare un modello di relazione fra variabile di risposta e determinanti eliminando l'effetto dei confondenti e controllando per i modificatori d'effetto. Pertanto è necessario stabilire in questa sezione, anche sulla base della letteratura, quali fattori vanno presi in considerazione come determinanti, quali rispondono alla definizione di confondente e quali possano essere potenziali modificatori d'effetto.

### Processi di misurazione:

vanno descritte le modalità di misura delle variabili di risposta ed esplicative; ad esempio per rilevazioni categoriche come per abitudini, stili di vita o preferenze vanno specificate le categorie che compongono la variabile (si veda tabella 2), mentre nel caso di variabili ordinali o a scala (lunga o corta) che come ad esempio quelle che esprimono i gradi di giudizio o di autosufficienza o di qualità della vita, devono essere descritte in dettaglio e specificato il numero delle domande che le costituiscono. Mentre nel caso di variabili i cui valori sono il risultato di un processo di misura vanno descritti quali strumenti verranno utilizzati e l'unità di misura adottate.



**Tabella 2.** - Esempi di tipi di variabili.

Variabile	Tipo/scala	Valori/categorie	Parametro
Quali bevande alcoliche ha consumato nell'ultimi 30 giorni?	Categorica/ nominale	<ul style="list-style-type: none"> <li>• Birra (esclusa analcolica)</li> <li>• Soft drink</li> <li>• Vino, porto, sherry</li> <li>• Liquori</li> <li>• Super alcolici</li> <li>• Non ho bevuto</li> </ul>	Frequenza relativa.
Grado di istruzione	Categorica/ ordinale	<ul style="list-style-type: none"> <li>• Elementare o Nessun titolo</li> <li>• Licenza media</li> <li>• Licenza superiore</li> <li>• Laurea triennale</li> <li>• Laurea magistrale</li> </ul>	Frequenza relativa
Esposizione al fumo (indiretto e diretto)	Categorica/ordinale	<ul style="list-style-type: none"> <li>• Non fumatore e non esposto fumo passivo</li> <li>• Non fumatore, esposto a fumo passivo</li> <li>• Fumatore modesto</li> <li>• Fumatore moderato</li> <li>• Forte fumatore</li> </ul>	Frequenza relativa
Lei è favorevole ad essere seguito da un infermiere?	Quasi-continua: Discreta/ scala corta (in generale tutte le scale da 3 a 5 item)	<ul style="list-style-type: none"> <li>• Decisamente favorevole,</li> <li>• Leggermente favorevole,</li> <li>• Né favorevole né contrario,</li> <li>• Leggermente contrario,</li> <li>• Decisamente contrario.</li> </ul>	Frequenza relativa o media
Numero eventi avversi, Numero figli, Numero visite MMG	Quasi-continua: Discreta/conteggi eventi poco numerosi	Numerico	Frequenza relativa o media
Auto sufficienza ECOG (0-5)	Quasi continua: Discreta/ scala lunga (in generale tutte le scale con più di 5 item)	Numerico	Media
Lesioni da decubito Scala di BRADEN (6-23)	Quasi-continua: Discreta/scala lunga	Numerico 6=alto rischio; 16=situazione a rischio 23=basso rischio	Media
Numero globuli rossi, Giornate di degenza Accessi PS	Quasi-continua: Discreta/conteggi eventi numerosi	Numerico	Media
Età; Altezza area volume Peso; Pressione diastolica	Misura Continua	Numerico	Media

**Parametri di misura:** sono le misure di sintesi per descrivere l'intensità del fenomeno collettivo. In generale si utilizza la media (aritmetica, geometrica, armonica) nel caso di variabili quantitative come le misure (ad es pressione sanguigna o temperatura corporea) oppure nel caso di variabili discrete come conteggi di eventi numerosi o variabili a scala lunga (come la scala di Braden) (Tabella 2). Mentre per variabili categoriche (nominali oppure ordinali) si utilizza la frequenza relativa per ogni categoria. Nel caso di variabili a scala corta (tabella 2) o che contano eventi poco numerosi, bisogna valutare i problemi di interpretazione e specificare se la media come indice di sintesi sia un buon rappresentante del fenomeno collettivo oppure è preferibile mostrare la frequenza relativa per singolo

valore. Nel caso si voglia indagare il grado di connessione fra due variabili, o si voglia misurare quanto il variare di una variabile implica il variare di un'altra, allora i parametri saranno: nel primo caso il coefficiente di correlazione, nel secondo caso il coefficiente di regressione.

**Le misure d'effetto e di connessione:** la misura d'effetto esprime il grado di influenza della variabile esplicativa sulla variabile di risposta; mentre la misura di connessione esprime il grado di variazione reciproca fra due variabili. Le misure di effetto esprimono una relazione gerarchica in termini di *relazione causale*: una variabile sarà la *causa* l'altra *l'effetto*, definibile come *relazione asimmetrica*, mentre le misure di connessione non mettono le

variabili in relazione gerarchica, infatti si può affermare che esprimono l'intensità di *relazione simmetrica*. Come abbiamo già accennato nel caso di studi esploratori-descrittivi saremo più interessati alla connessione fra variabili, mentre negli studi analitico-valutativi l'interesse si sposta sulla relazione *causa-effetto*. Quindi l'intento dello studio che guiderà la scelta del disegno di studio, assieme al tipo delle variabili

determineranno la scelta delle misure d'effetto e di connessione. Nella tabella 3 vengono illustrate alcune misure d'effetto e di connessione a seconda del tipo di relazione indagata e per tipo di variabili coinvolte. Gli esempi riguardano modelli a due variabili (bi-variati).

Tabella 3. - Misure d'effetto o di connessione.

Variabile esplicativa	Variabile di risposta	Parametri	Tipo relazione che si intende indagare	Misura
Dicotomica (es. fumatore/ non-fumatore)	Continua (o quasi continua) (pressione diastolica)	Media della pressione tra i fumatori, Media della pressione tra i non-fumatori	Asimmetrica	Differenza di medie,
Dicotomica (es fumatore/ non-fumatore)	Dicotomica: diagnosi di asma bronchiale (si/no)	Frequenza delle diagnosi fra i fumatori (rischio di asma fra i fumatori) Frequenza delle diagnosi fra i non fumatori (rischio di asma fra i non-fumatori)	Asimmetrica	Differenza di rischi; oppure Rapporto di rischi.
			Simmetrica	OddsRatio
Politomica non ordinale (ad es. mansione lavorativa)	Continua (o quasi continua) ad es concentrazione piombo ematico	Media dei valori per tipo di mansione	Asimmetrica	Confronto fra medie
Politomica Ordinale (esposizione al fumo)	Continua (quasi Continua) Funzionalità polmonare (Flusso Espiratorio Forzato)	Media dei valori per livello di esposizione	Asimmetrica	Confronto fra medie, trend lineare,
Politomica ordinale	Politomica ordinale	Coefficiente correlazione fra ranghi	Simmetrica	Coefficiente correlazione fra ranghi
Politomica non ordinale	Politomica non ordinale	Connessione Chi-quadrato Phi o V di Cramer	Simmetrica	Connessione Chi-quadrato Phi o V di Cramer
Politomica Ordinale (esposizione al fumo)	dicotomica (diagnosi sofferto di bronchite negli ultimi 12 mesi) (si /no)	Frequenza numero casi di bronchite nei vari livelli della variabile esplicativa	Asimmetrica	Confronto di frequenze relative, trend lineare
Continua (quasi continua)	Continua (quasi continua)	Coefficiente di regressione	Asimmetrica	Coefficiente di regressione
		Coefficiente di correlazione lineare o correlazione fra ranghi	Simmetrica	Coefficiente di correlazione lineare o correlazione fra ranghi

**Popolazione osservabile o accessibile:** sono le unità della popolazione obiettivo effettivamente raggiungibili, quindi va definito il criterio utilizzato per raggiungerle (ad es tutti i pazienti con disabilità intellettiva presenti nel registro). In altri termini, vanno descritti e motivati sia i criteri di inclusione nella popolazione, sia i criteri di esclusione; questi ultimi saranno cruciali per determinare se i risultati saranno generalizzabili alla popolazione obiettivo definita nella fase precedente. Ad

esempio i criteri di esclusione potrebbero riguardare periodi di tempo, o presenza di patologie concomitanti. I motivi di esclusione vanno dettagliati e giustificati.

**Protocollo di campionamento:** L'obiettivo delle procedure di campionamento è garantire che il campione (la base empirica dello studio) sia rappresentativo della popolazione osservabile, ovvero che ne possiede tutte le caratteristiche note ed ignote, così che le

eventuali differenze siano da attribuirsi solo al caso. Inoltre deve chiarire sia i metodi di reclutamento dei soggetti dalla popolazione osservabile, sia il criterio di calcolo della dimensione campionaria.

**Protocollo di raccolta dati:** è il documento che esplicita l'intero processo di raccolta dati e dà indicazioni operative su quanto descritto nei paragrafi precedenti. Deve contenere anche i metodi di gestione Informativa/informatica dei dati.

Tale documento deve riportare le variabili utili all'indagine, il tipo di valori che si intendono raccogliere e le modalità con i quali tali dati sono raccolti.

I dati possono essere raccolti *ad hoc* o provenire da fonti preesistenti; nel primo caso nei paragrafi precedenti abbiamo chiarito che bisogna specificare le unità di osservazione e le loro caratteristiche (variabili), il protocollo deve riportare gli schemi utilizzati per le raccolte dei dati (schede, questionari ecc.) e le modalità di somministrazione. Nel caso si utilizzino dati da fonti preesistenti (ad es schede dimissione ospedaliera, dati di laboratorio, registri di malattia), ciò che è stato specificato nei paragrafi precedenti rimane valido.

Il protocollo deve specificare la natura e la struttura di tutte le fonti dati, così che il modello della realtà, che scaturisce dall'analisi del problema, possa essere trasformato in una struttura dati (fase 3) che consenta l'analisi statistica (fase 4). Il protocollo di raccolta dati rappresenta un documento fondamentale per garantire la registrazione di tutte le informazioni pertinenti; nel caso siano coinvolti più operatori il protocollo garantisce l'uniformità di comportamento, infine fornisce la documentazione sulla struttura dei dati al fine di agevolare la realizzazione tecnica del data base e il suo mantenimento per la durata dello studio.

### FASE 3: Dati – Struttura dei dati

In questa fase l'obiettivo è pianificare la trasformazione del modello concepito nel disegno di studio in una struttura dati che consenta la registrazione dei dati raccolti *ad hoc*

o la gestione delle strutture dati preesistenti. L'obiettivo nel creare una struttura dati è quello di consentire le valutazioni di qualità dei dati e le analisi statistiche necessarie a rispondere al quesito. Sottolineiamo che queste tre fasi *Piano-Dati-Analisi* sono profondamente interdipendenti, pertanto è possibile che la pianificazione dell'analisi statistica possa costringere a rivedere la struttura dei dati definita al passo precedente.

**Struttura dei dati:** nel caso si debbano raccogliere dati *ad hoc*, la struttura dati più semplice è costituita da un'unica tabella, le cui righe sono dette record e rappresentano le unità di osservazione, mentre le colonne rappresentano le variabili, ovvero le caratteristiche o attributi di ogni unità di osservazione. Ogni colonna/variabile è dotata di un Nome (nome della variabile) e di un Valore (assunto da quella specifica entità/record), inoltre ogni variabile può assumere valori di un solo tipo (numerico, alfanumerico, data). I possibili valori che la specifica variabile può assumere devono essere definiti in precedenza come illustrato nel paragrafo "processi di misurazione" tabella 2. Ogni struttura dati elementare (tabella) deve essere descritta tramite un documento chiamato **tracciato record** che elenca campi e i loro formati (ovvero i valori ammessi) (tabella 4). Nel caso di ricerche semplici, come può essere una valutazione di gradimento del servizio da parte dell'utenza, un database costituito da una sola tabella è sufficiente, mentre nel caso di ricerche più complesse che richiedono registrazioni dati su più livelli; ad esempio le informazioni dei dati anagrafici del paziente (livello paziente) potrebbero richiedere una struttura a parte, separata dalla tabella che registra le caratteristiche del tipo di visita (livello visita) e così via; in questi casi è necessario concepire una struttura dati (database) che consente di gestire simultaneamente più tabelle elementari, collegate fra loro da specifiche chiavi di collegamento; fra queste strutture più complesse la più usata è quella detta *relazionale* che rappresenta il data base come insieme di tabelle collegate fra loro.

**Tabella 4.-** Esempio tracciato record.

Varabile	Tipo	Lunghezza	Valori ammessi
Identificativo del soggetto	Numerico	6	
Età-anni compiuti	Numerico	2	
Tipo di disabilità	Numerica	2	0=nessuna 1=lieve 2=moderata 3=grave 4=profonda 99=non rilevata
Data diagnosi	Data	109	gg/mm/aaaa
Presenza di caregiver	Numerico	1	0=no 1=si
Numero richiesta di assistenza sanitaria negli ultimi 12 mesi	Numerico	3	

**Vincoli di integrità dei dati:** Contestualmente alla costruzione della base dati vanno chiariti i meccanismi (o le procedure) di controllo che ne garantiscono la coerenza interna e che saranno parte integrante del sistema informatico, come ad esempio prevenire l'inserimento di valori non previsti per le variabili o valori non congruenti con altre informazioni presenti per la stessa unità statistica, ad esempio genere maschile e patologie esclusivamente femminili. Inoltre le procedure di raccolta dati dovrebbe prevenire la circostanza in cui un valore per una data variabile non viene registrato pur essendo presente (dato mancante); in questi casi il sistema deve poter distinguere la circostanza di *dato mancante* dalla situazione in cui l'informazione non viene raccolta poiché non presente per quell'unità di osservazione (*zero strutturale*), ciò tipicamente accade quando la rilevazione non risulta applicabile a quella unità statistica.

**Data Screening e Pulizia dei dati:** in generale, e soprattutto nei casi in cui i dati risiedono su archivi preesistenti, è sempre opportuno prevedere le procedure di controllo dei dati genericamente indicate come *data screening*. Tali procedure aiutano ad utilizzare in modo appropriato una base di dati; puntano ad isolare le particolarità dei dati, come: valori mancanti o zero-strutturali, oppure valori errati o anomali ad es troppo grandi o troppo piccoli (*outlier*), record duplicati o record vuoti, cioè con pochissime variabili valorizzate, oppure i gruppi di casi insoliti, ovvero unità statistiche che presentano combinazioni insolite di informazioni che possono nascondere possibili errori di inserimento; ad esempio: un soggetto di 15 anni

con un reddito di 90000 euro l'anno (non è impossibile, ma certamente è insolito). Inoltre, il data screening permette di stabilire le caratteristiche distributive dei dati in previsione dell'applicazione dei metodi di analisi statistica, ovvero se i dati presentano le caratteristiche necessarie per poter applicare i metodi della statistica inferenziale, consentendo di stabilire quali adattamenti applicare.

#### FASE 4: Analisi

Lo scopo dell'Analisi è quello di utilizzare sistematicamente i dati e le informazioni raccolti secondo il disegno di studio (Piano) per affrontare le domande formulate nella fase di definizione del problema. Il tipo di Analisi statistica dipende dalla complessità del problema, dal disegno di studio, dal tipo di dati e, non ultimo, anche dai possibili lettori (MacKay e Oldford, 2000).

Il primo passo dell'analisi statistica di tipo descrittivo e si serve di indici statistici e di grafici per descrivere le caratteristiche del campione in studio. Nel caso più volte citato del confronto fra due gruppi definiti dai due livelli del determinante (esposto/non-esposto) è importante rendere una descrizione per ogni gruppo al fine di evidenziare eventuali differenze; ad esempio la proporzione di maschi e femmine, l'età media o la proporzione per fasce di età, o per categorie definite per le condizioni di salute o il punteggio medio di un indice di auto sufficienza e così via.

Il secondo passo dell'analisi è quello inferenziale, che consiste nel sottoporre a verifica l'ipotesi in studio o, se vogliamo, il modello ipotizzato nel disegno di studio. In genere questa fase dell'analisi richiede di

confrontare i valori della variabile di risposta (outcome) tra i vari livelli del determinante.

### FASE 5: Conclusioni

Lo scopo di quest'ultima fase è quello di riferire i risultati dello studio nel linguaggio del Problema. Il gergo statistico dovrebbe essere evitato. Inoltre, offre l'opportunità di discutere i punti di forza e di debolezza del Piano, dei Dati e dell'Analisi soprattutto per quanto riguarda i possibili errori che potrebbero essersi verificati (MacKay e Oldford, 2000).

In molti lavori la parte conclusiva si apre con la **discussione dei risultati**: dove si devono commentare e dare la propria interpretazione dei risultati ottenuti, evidenziando se i risultati sono in accordo o in disaccordo con l'ipotesi iniziale oppure sono inconclusivi (non statisticamente significativi). Vanno discusse le similarità e differenze dei risultati dello studio con quelli ottenuti da altri autori sul medesimo quesito di ricerca. La parte conclusiva si chiude con **le conclusioni dello studio**: dove vanno ulteriormente condensate le conoscenze più importanti acquisite tramite il lavoro, discussi i potenziali limiti dello studio ed evidenziata l'importanza dei risultati ottenuti e chiarire le implicazioni dei risultati per la pratica e per la ricerca futura.

### DISCUSSIONE

L'intento principale di questo contributo è quello di presentare il ciclo PPDAC come modello di unitario di pianificazione delle ricerche statistiche, già impiegato per l'educazione al metodo statistico (MacKay e Oldford, 2000), poiché ha il pregio di armonizzare in un unico schema parti normalmente trascurate, come gli aspetti di gestione informatica dei dati e quelli sulle analisi statistiche. Inoltre la struttura ciclica del PPDAC, che prevede un approccio ricorsivo fra le fasi, si rivela una buona guida alla progettazione delle ricerche.

La moderna ricerca scientifica raramente può essere condotta da una persona sola. Per eseguire una buona ricerca o un buon progetto di ricerca, ancor prima di pensare a come raccogliere i dati, occorre costruire una buona squadra che deve riunire diverse competenze e diversi punti di vista sul quesito che si vorrebbe affrontare. Oltre alle competenze di merito, occorrono competenze di metodo e tecnologiche e, come abbiamo accennato, devono essere inclusi anche i possibili portatori di interesse: professionisti e, nel caso della ricerca clinica, i pazienti. In ogni caso è bene che tutti i partecipanti alla ricerca siano consapevoli dei criteri di progettazione e dei metodi di conduzione della ricerca.

### Conflitto di interessi

Tutti gli autori dichiarano l'assenza di conflitto di interessi. Tutti gli autori dichiarano di aver contribuito alla realizzazione del manoscritto e ne approvano la pubblicazione.

### Finanziamenti

Gli autori dichiarano di non aver ottenuto alcun finanziamento e l'assenza di sponsor economici.

### BIBLIOGRAFIA

- AA.VV.(2021). *Florence Nightingale e l'Italia due secoli di arte e scienza infermieristica*. FNOPI. consultato il 26 maggio 2022 da: <https://www.fnopi.it/wp-content/uploads/2021/05/NIGHTINGALE.pdf>.
- Al-Jundi A., Sakka S. (2016). *Protocol Writing in Clinical Research*. J Clin Diagn Res. 2016 Nov; 10(11): ZE10–ZE13.
- Bando K, Sato T (2015). *Did you write a protocol before starting your project?* Gen Thorac Cardiovasc Surg. 2015 Feb;63(2):71-7.
- Betts H.J., Wright G. (2020). *200 years since the birth of nursing informatics?* Stud Health Technol Inform; 272:28-32.
- Bur J, Missen K, Cooper S. (2021) *The impact of intellectual disability nurse specialists in the United Kingdom and Eire Ireland: An integrative review*. Nurs Open. Sep;8(5):2018-2024.
- Copi I.M., Cohen C. (1999). *Introduzione alla logica*. Prentice Hall International, Società editrice il Mulino. ISBN 88-15-07217-9.
- de Smith M.J. (2018). *STATSREF: Statistical Analysis Handbook - a web-based statistics resource* *Statistical Analysis Handbook* consultato il 30 maggio 2022 da: [https://www.statsref.com/HTML/?statistics\\_\\_statistical\\_analys.html](https://www.statsref.com/HTML/?statistics__statistical_analys.html).
- Hassmiller S. (2021) *Florence Nightingale as an Applied Statistician*. AJN, American Journal of Nursing: - Volume 121 - Issue 9 - p 71.
- Higgins J.P.T., Thomas J., Chandler J., Cumpston M., Li T., Page M.J., Welch V.A. (editors) (2019). *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd Edition. Chichester (UK): John Wiley & Sons.
- Kopf E.W. (1916). *Florence nightingale as statistician*. Q Publ Am Stat Assoc. 15(116):388–404.
- Lazzari G., Salvini L., Patella S., Ausili D., Di Mauro S. (2015). *La strutturazione del quesito clinico per reperire le prove di efficacia. una revisione della letteratura*. Rivista L'Infermiere N° 6.
- Levy M.L., Robb M., Allen J., Doherty C., Bland J.M. and Winter J.D. (2000) *A randomized controlled evaluation of specialist nurse education following*



accident and emergency department attendance for acute asthma. *Respiratory Medicine* 94, 900-908.

MacKay R.J., Oldford R.W. (2000). *Scientific Method, Statistical Method and the Speed of Light*. *Statist. Sci.*, 15, 3, 254-278 consultato il 26 maggio 2022 da: <https://projecteuclid.org/euclid.ss/1009212817>.

Nuttall P. (1983). *The passionate statistician (Florence Nightingale)*. *Nursing Time* 45september 28.

O'Brien K. Wright J. (2022). *How to write a protocol*. *Journal of Orthodontics*, vol 29, 58-61.

Pishgooie A.H., Atashzadeh-Shoorideh F., Falcó-Pegueroles A., Lotfi Z. *Correlation between nursing managers' leadership styles and nurses' job stress and anticipated turnover*. *J Nurs Manag.* 2019;27:527–534.

Polit D.F., Beck C.T. (2014). *Fondamenti di ricerca Infermieristica*. Ed McGraw Hill.

Rothman J.K. (2007). *Epidemiologia*. Casa Editrice Idelson-Gnocchi.

Rice V.H., Heath L., Livingstone-Banks J., Hartmann-Boyce J. (2017) *Nursing interventions for smoking*

cessation. *Cochrane Database of Systematic Reviews*, Issue 12. Art.

Treccani (2022a) consultato il 26 maggio 2022 da: <https://www.treccani.it/vocabolario/inferire/>.

Treccani (2022b) consultato il 26 maggio 2022 da: <https://www.treccani.it/vocabolario/inferenza/>.

Treccani (2022c) consultato il 26 maggio 2022 da: <https://www.treccani.it/vocabolario/deduzione/>.

Treccani (2022d) consultato il 26 maggio 2022 da: <https://www.treccani.it/vocabolario/induzione/>.

Wild C.J., Pfannkuch M.. (1999). *Statistical Thinking in Empirical Enquiry*. *International Statistical Review* 199. 67,3,223-265. consultato il 26 maggio 2022 da: <https://iase-web.org/documents/intstatreview/99.Wild.Pfannkuch.pdf>.

Cochrane Work (2022). *Registering a new title*. consultato il 07 giugno 2022 da: <https://work.cochrane.org/registering-new-title>.