



<b>Denominazione</b>	<b>Corso integrato: Statistica ed analisi dei dati</b> <b>Integrated course: Statistics and data analysis</b>
<b>Moduli componenti</b>	Metodologia epidemiologica Statistica medica avanzata Sistemi informativi in medicina Tecnologie AI e Big data
<b>Settore scientifico-disciplinare</b>	Metodologia epidemiologica MEDS-24/B Statistica medica avanzata MEDS-24/A Sistemi informativi in medicina IBIO-01/A Tecnologie AI e Big data IINF-05/A
<b>Anno di corso e semestre di erogazione</b>	<b>III° ANNO-I° SEMESTRE</b>
<b>Lingua di insegnamento</b>	Italiano
<b>Carico didattico in crediti formativi universitari</b>	<b>6 CFU totali</b> , distribuiti come segue: 2 CFU (Metodologia epidemiologica) 1 CFU (Statistica medica avanzata) 1 CFU (Sistemi informativi in medicina) 2 CFU (Tecnologie AI e Big data)
<b>Numero di ore di attività didattica frontale</b>	20 ore Didattica Erogativa (DE) + 5 Didattica Interattiva (DI) (Metodologia epidemiologica) 10 ore Didattica Erogativa (DE) + 2.5 Didattica Interattiva (DI) (Statistica medica avanzata) 10 ore Didattica Erogativa (DE) + 2.5 Didattica Interattiva (DI) (Sistemi informativi in medicina) 20 ore Didattica Erogativa (DE) + 5 Didattica Interattiva (DI) (Tecnologie AI e Big data)
<b>Docenti</b>	Licia Iacoviello (Metodologia epidemiologica) [ <b>Coordinatore del CI</b> ] Alessandro Gialluisi (Statistica medica avanzata) Alessandro De Laurentiis (Sistemi informativi in medicina) Antonietta Pepe (Tecnologie AI e Big data)
<b>Obiettivi formativi</b>	Gli obiettivi formativi del corso sono fornire gli elementi essenziali per poter eseguire valutazioni in ambito epidemiologico con metodi e strumenti appropriati rispetto all'obiettivo di tutela della salute pubblica e di preparare lo studente a conoscere i metodi fondamentali per osservare e promuovere la salute del singolo e della comunità nonché quelle relative ai compiti del medico in tale campo, in particolare per gli strumenti della prevenzione primaria, secondaria e terziaria. Inoltre, il corso, con il modulo di Statistica medica avanzata, si propone di introdurre lo studente ai principi elementari della ricerca in medicina quantitativa, dove l'oggetto di studio non è un singolo individuo ma un collettivo. In particolare lo studente imparerà a: costruire e interpretare una tabella di frequenza a partire dai dati individuali; rappresentare adeguatamente in forma grafica i dati relativi a un fenomeno biologico; saper calcolare gli intervalli di riferimento (di normalità) di una variabile biologica; saper valutare il rischio relativo associato a una data esposizione e la sensibilità, specificità e valore predittivo di un test diagnostico; saper effettuare un test per il confronto tra due medie o delle proporzioni; misurare l'effetto di un trattamento e interpretare il suo intervallo di confidenza; stimare la relazione tra due variabili biologiche mediante un modello di regressione lineare. Obiettivo del corso è anche fornire i principi base della pianificazione degli studi osservazionali, interpretazione delle misure di occorrenza di malattia e di associazione tra fattori di rischio e occorrenza di malattia e della analisi delle sperimentazioni cliniche e del concetto di inferenza statistica. Nel corso lo studente acquisirà anche conoscenze sugli attuali sistemi disponibili per il trasferimento delle informazioni e competenze nella gestione dei dati e sulla loro elaborazione avanzata con risorse intelligenza artificiale e machine learning applicate alla ricerca clinica.
<b>Risultati di apprendimento specifici</b>	<i>Conoscenze e comprensione.</i> Lo studente apprenderà i principi elementari della ricerca in medicina quantitativa, dove l'oggetto di studio non è un singolo individuo ma un collettivo. In particolare lo studente dovrà: <ul style="list-style-type: none"><li>• comprendere i concetti chiave dell'epidemiologia e la loro applicazione nella medicina, compresa la distribuzione delle malattie, i determinanti della salute e l'utilizzo delle misure epidemiologiche;</li><li>• acquisire le competenze per valutare i dati epidemiologici e interpretare i risultati degli studi;</li></ul>



- riconoscere e valutare i diversi tipi di disegni di studi epidemiologici;
- analizzare le distribuzioni delle malattie nella popolazione e identificare i fattori di rischio/prevenzione associati con esposizioni genetiche e ambientali (in particolare riguardo all'alimentazione e stili di vita e alle condizioni socioeconomiche);
- comprendere le misure di frequenza delle malattie e le misure di associazione tra fattori di rischio e malattie;
- valutare la validità e l'affidabilità dei dati epidemiologici;
- conoscere i principali metodi di statistica descrittiva ed inferenziale, i più comuni metodi per studiare relazioni tra variabili, in particolare con outcome di interesse clinico;
- conoscere e saper applicare i principali metodi di statistica descrittiva ed inferenziale, su una o più variabili nelle diverse tipologie di studio osservate, per verificare le ipotesi di ricerca mediante excel e software statistici disponibili online;
- comprendere gli approcci più moderni all'utilizzo dei dati sanitari: i big-data e l'intelligenza artificiale;
- conoscere i principali algoritmi di machine learning applicati a dati biomedici, avere una comprensione di base del loro funzionamento e della terminologia adeguata;
- comprendere i differenti ambiti di applicazione dell'analisi di Big Data e delle tecniche di AI in ambito sanitario, e le loro potenzialità;
- apprendere i principali concetti relativi ai sistemi informativi in ambito clinico e sanitario.

*Capacità di applicare conoscenze e comprensione.*

Al termine delle lezioni dell'insegnamento lo studente dovrà essere in grado di:

- descrivere la differenza fra epidemiologia descrittiva ed analitica;
- spiegare le differenze fra i concetti di prevalenza e incidenza;
- spiegare il concetto di rischio ed effettuare semplici calcoli del rischio relativo, assoluto e attribuibile;
- interpretare i tassi standardizzati con il metodo diretto e indiretto;
- descrivere le caratteristiche, i vantaggi e i limiti degli studi epidemiologici;
- saper leggere in modo critico la letteratura scientifica e poter condurre un'analisi statistica, interpretandone correttamente i risultati.
- identificare un caso tipico di applicazione di tecniche di AI in ambito sanitario ed i casi ideali di utilizzo di tali applicazioni.
- descrivere le principali caratteristiche dei sistemi informativi sanitari e clinici e la loro correlazione rispetto ai processi di diagnosi e cura.

*Abilità comunicative*

Al termine delle lezioni dell'insegnamento lo studente dovrà:

- saper utilizzare il linguaggio adeguato ad illustrare gli strumenti epidemiologici acquisiti;
- saper utilizzare i principali disegni di studio epidemiologico osservazionale o sperimentale;
- costruire e interpretare una tabella di frequenza a partire dai dati individuali;
- rappresentare adeguatamente in forma grafica i dati relativi a un fenomeno biologico;
- saper calcolare gli intervalli di riferimento (di normalità) di una variabile biologica;
- saper valutare il rischio relativo associato a una data esposizione e la sensibilità, specificità e valore predittivo di un test diagnostico;
- saper effettuare un test per il confronto tra due medie o due proporzioni;
- saper misurare l'effetto di un trattamento e interpretare il suo intervallo di confidenza;
- saper stimare la relazione tra due variabili biologiche mediante un modello di regressione lineare utilizzando strumenti di Intelligenza Artificiale;
- saper interpretare e presentare correttamente i risultati di un'analisi statistica descrittiva ed inferenziale, con particolare riguardo ai metodi più comunemente utilizzati negli studi epidemiologici;
- saper interpretare e spiegare il flusso tipico di analisi di un algoritmo di machine learning applicato a dati biomedici, i principali risultati;
- saper identificare e spiegare le potenziali applicazioni delle tecniche di AI in campo sanitario;
- saper individuare quali siano i sistemi informativi di cui necessita per lo svolgimento della pratica medica, sia in ambito ospedaliero che territoriale e descriverne i principali vantaggi.

*Autonomia di giudizio e pensiero critico.*

Lo studente maturerà la capacità di analizzare criticamente i risultati degli studi epidemiologici osservazionali e sperimentali di valutazione dell'efficacia degli interventi medici.



<b>Programma</b>	<p><b>Metodologia epidemiologica</b></p> <ul style="list-style-type: none"><li>● Introduzione all'epidemiologia medica: definizioni, scopo e principi fondamentali.</li><li>● Le misure in epidemiologia: Misure di frequenza: prevalenza, incidenza, tassi di incidenza, standardizzazione dei tassi.</li><li>● I principali studi epidemiologici osservazionali: ecologici, trasversali, caso controllo, di coorte.</li><li>● Gli studi epidemiologici sperimentali: Studi clinici di fase I, II, III, IV</li><li>● Gli studi epidemiologici secondari: Ricerca bibliografica, revisione narrativa, revisione sistematica, principi di metanalisi.</li><li>● Le principali fonti di dati per gli studi epidemiologici.</li><li>● Le misure di associazione: rischio relativo, odds ratio, rischio assoluto e rischio attribuibile.</li><li>● Valutazione della validità e dell'affidabilità degli studi epidemiologici.<ul style="list-style-type: none"><li>○ Bias e confondimento: bias di selezione, di informazione e confondimento.</li><li>○ Metodi di controllo del confondimento: scelta del campione, raccolta dati, randomizzazione, restrizione, appaiamento, analisi stratificate e multivariate.</li><li>○ Interazione e mediazione di azione.</li></ul></li><li>● Il concetto di rischio e la predizione del rischio.</li><li>● Epidemiologia dei fattori di rischio e il concetto di "common soil".</li><li>● Epidemiologia sociale e disuguaglianze nella salute. Il concetto di "Global Health".</li><li>● Epidemiologia della nutrizione e della dieta mediterranea.</li></ul> <p><b>Statistica medica avanzata</b></p> <p>Statistica descrittiva:</p> <ul style="list-style-type: none"><li>● Scale di misura. Tipi di variabili.</li><li>● Distribuzioni di frequenza. Tabelle. Grafici.</li><li>● Misure di posizione (media, mediana, moda, percentili e quartili) e di dispersione (range min-max, interquartile, varianza, deviazione standard, coefficiente di variazione).</li></ul> <p>Statistica inferenziale:</p> <ul style="list-style-type: none"><li>● Concetto di Stima: puntuale ed intervallare.</li><li>● Test di verifica delle ipotesi:<ul style="list-style-type: none"><li>○ Test su una media, su due medie (campioni indipendenti e appaiati).</li><li>○ Test non parametrici (campioni indipendenti e appaiati).</li><li>○ Verifica di ipotesi per il confronto tra più medie e confronti multipli.</li><li>○ Verifica di ipotesi su una proporzione e per il confronto tra due proporzioni.</li><li>○ Tabelle di contingenza. Il test <math>\chi^2</math>. Misure di associazione (Odds Ratio e Hazard Ratio).</li><li>○ Concetto di potenza statistica e numerosità campionaria.</li></ul></li><li>● Correlazione parametrica e non parametrica (cenni).</li><li>● Regressione lineare, logistica e multipla.</li><li>● Analisi della sopravvivenza: curve di Kaplan-Meier e regressione di Cox.</li></ul> <p>Algoritmi di classificazione, test diagnostici e di screening:</p> <ul style="list-style-type: none"><li>● Parametri di valutazione (sensibilità, specificità, valore predittivo positivo e negativo).</li><li>● Curva ROC.</li><li>● Utilizzo dei big data e dell'intelligenza artificiale (cenni).</li></ul> <p>Esercitazioni con tools statistici.</p> <p><b>Sistemi informativi in medicina</b></p> <ul style="list-style-type: none"><li>● Sistemi informativi sanitari e clinici</li><li>● Informatizzazione dei processi di diagnosi e cura</li><li>● Cartella clinica elettronica e dossier sanitario.</li><li>● Sistemi informativi territoriali</li><li>● Telemedicina</li><li>● Privacy e trattamento dei dati sanitari.</li></ul> <p><b>Tecnologie AI e Big data</b></p> <ul style="list-style-type: none"><li>● Big data e AI: definizioni ed applicazioni in Healthcare</li><li>● Estrazione, integrazione e trasformazione dei dati: fondamenti teorici ed esercitazioni pratiche.</li><li>● Fondamenti di Information Visualization per la creazione di visualizzazioni di dati efficaci.</li></ul>
------------------	--



	<ul style="list-style-type: none"><li>• Flussi automatizzati di apprendimento automatico per Descriptive e Predictive Analytics: esempi di regressione, classificazione, e clusterizzazione di dati bio-medicali.</li><li>• Cenni sui concetti di training, testing, tuning e trasparenza di un modello.</li></ul>
<b>Tipologie di attività didattiche previste e relative modalità di svolgimento</b>	<p>Il corso sarà tenuto attraverso una combinazione di lezioni frontali ed esercitazioni pratiche, con la formazione di gruppi di lavoro sul disegno e la conduzione di studi epidemiologici, la valutazione dei dati epidemiologici e la lettura critica degli studi. Gli studenti saranno incoraggiati a partecipare attivamente alle lezioni e a porre domande per una migliore comprensione dei concetti. Verranno presentati casi reali e scenari per applicare concetti e metodi epidemiologici.</p> <p>Gli studenti verranno coinvolti in esercitazioni pratiche, analisi dei dati e interpretazione dei risultati epidemiologici. Particolare attenzione verrà posta alla promozione del pensiero critico, della risoluzione dei problemi e delle competenze decisionali.</p> <p>Verranno organizzati seminari con relatori ospiti come epidemiologi esperti per condividere la loro esperienza e le loro conoscenze e tavole rotonde su argomenti attuali e controversi in epidemiologia con lo scopo di favorire opportunità di networking e collaborazione tra studenti e professionisti nel campo dell'epidemiologia e della statistica.</p> <p>È previsto l'utilizzo di tecnologie digitali per l'erogazione delle lezioni e delle esercitazioni.</p>
<b>Metodi e criteri di valutazione dell'apprendimento</b>	<p>La valutazione dell'apprendimento consiste nello svolgimento di una prova scritta contenente quesiti teorici e pratici relativi ai contenuti del corso, proposti attraverso domande a risposta multipla.</p> <p>Sarà prevista inoltre una prova orale facoltativa. In questo caso, la prova scritta concorrerà alla composizione del voto finale, nella misura del 70%. Il restante 30% della valutazione si baserà sul colloquio orale finale. La valutazione del colloquio orale è espressa in trentesimi e terrà conto della proprietà di linguaggio, della capacità argomentativa, di analisi critica e di ragionamento.</p>
<b>Criteri di misurazione dell'apprendimento e di attribuzione del voto finale</b>	<p>La valutazione dell'apprendimento prevede l'attribuzione di un voto finale espresso in trentesimi. Il voto finale sarà determinato attraverso l'esame scritto che si compone di 32 quesiti a risposta multipla (1 punto per ciascun quesito).</p> <p>La concessione della lode sarà valutata per i soli studenti che abbiano raggiunto la valutazione complessiva di 32/32.</p> <p>La prova di esame qualora integrata dall'esame orale sarà valutata secondo i seguenti criteri:</p> <p><b>NON IDONEO:</b> importanti carenze e/o inaccuratezza nella conoscenza e comprensione degli argomenti; limitate capacità di analisi e sintesi, frequenti generalizzazioni.</p> <p><b>IDONEO:</b></p> <ul style="list-style-type: none"><li>• <b>18-20:</b> conoscenza e comprensione degli argomenti appena sufficiente con possibili imperfezioni; capacità di analisi, sintesi e autonomia di giudizio sufficienti.</li><li>• <b>21-23:</b> Conoscenza e comprensione degli argomenti routinaria; capacità di analisi e sintesi corrette con argomentazione logica coerente.</li><li>• <b>24-26:</b> Discreta conoscenza e comprensione degli argomenti; buone capacità di analisi e sintesi con argomentazioni espresse in modo rigoroso.</li><li>• <b>27-29:</b> Conoscenza e comprensione degli argomenti completa; notevoli capacità di analisi, sintesi. Buona autonomia di giudizio.</li><li>• <b>30-30L:</b> Ottimo livello di conoscenza e comprensione degli argomenti. Notevoli capacità di analisi e di sintesi e di autonomia di giudizio. Argomentazioni espresse in modo originale.</li></ul>
<b>Propedeuticità</b>	<p>Per essere ammesso a sostenere l'esame del corso integrato lo studente dovrà aver sostenuto tutti gli esami del primo anno.</p> <p>E' desiderabile una buona conoscenza della lingua inglese.</p>

**Materiale didattico  
utilizzato e materiale  
didattico consigliato**

- Slides, dispense e materiale supplementare forniti dal docente.
- Carlo Signorelli, Anna Odone. *Elementi di metodologia epidemiologica*. VIII edizione. Società Editrice Universo (ISBN: 978-88-6515-21-64)
- M. Bland: *Statistica Medica*. APOGEO (ISBN: 978-8891629739) o in alternativa W.W. Daniel: *Biostatistica*. EdiSES (ISBN: 9788833190419)
- Andrea De Mauro. *Data Analytics per tutti*. Apogeo (ISBN: 978-8850335947)
- Alberto Rosotti. *Informatica Medica 2/ed.* Mc Graw Hill (ISBN: 978-8838699894)