

# TESTO VIDEOTUTORIAL DI INTRODUZIONE A PUBMED

A cura della biblioteca medica Pinali, Università degli Studi di Padova.

## *Pubmed – parte 1*

[HOME PAGE DI PUBMED]

Cos'è Pubmed: è una banca dati bibliografica, archivio di citazioni di area medica e aree vicine, è curata da una sezione della National Library of Medicine di Bethesda, NCBI il National Center for Biotechnology Information, e messa a disposizione gratuitamente dal 1996.

Contiene circa 23 milioni di citazioni bibliografiche tratte da più di 5600 riviste scientifiche internazionali, e da un numero ridotto di citazioni da monografie.

Le citazioni sono in inglese indipendentemente dalla lingua in cui è stato scritto l'articolo originale e per l'82% circa accompagnate da riassunto o abstract (se fornito dall'autore in inglese).

Sono presenti numerosi collegamenti, p.es ad altre risorse di biologia molecolare, genetica etc., ad articoli a testo integrale, a siti di editori, etc.

Contiene tutti i records Medline oltre a citazioni in fasi diverse di elaborazione o non destinate affatto ad alimentare Medline. Accennerò a breve alla differenza tra Pubmed e a Medline.

-

### **Descrizione di Home page di Pubmed**

Pubmed è una banca dati molto articolata e non sempre intuitiva. Cercherò di dare gli elementi essenziali della struttura e dell'utilizzo.

A sinistra abbiamo diversi link a istruzioni, domande frequenti, descrizione di funzioni, manuali, tutorial, aggiornamenti. Gli aggiornamenti sono descritti in **New and Noteworthy**. E' lì che bisogna andare a vedere se interviene qualche novità.

Infatti i lavori di aggiornamento sono continui. Tra gli aggiornamenti cadenzati c'è quello del vocabolario Mesh: ogni anno vengono introdotti termini nuovi e sostituiti quelli obsoleti. Del vocabolario Mesh parleremo ampiamente.

Al centro: l'unico link che apriremo è quello al **Single citation matcher**, mentre a destra guarderemo bene il **Mesh database**.

Un link interessante è quello a **Clinical Queries** nella colonna centrale dei titoli che permette di compiere ricerche con strategie predeterminate che difficilmente un utente non esperto riuscirebbe a costruire.

Cominciamo dal **Single Citation matcher**.

Dà l'accesso a un box che facilita la ricerca delle citazioni attraverso campi predisposti. E' molto intuitivo: ma è meglio non esagerare con i dettagli nella compilazione: si inseriscono le informazioni che si hanno (non troppe né troppo poche) cambiando strategia e compilando campi diversi in caso di insuccesso.

Es. cerchiamo il lavoro di Bianchini G su Cancer Research del 2010

*Prognostic and therapeutic implications of distinct kinase expression patterns in different subtypes of breast cancer.*

**N.B. Faccio la ricerca due volte mostrando strategie leggermente diverse**

**1<sup>a</sup> prova**

Cerco **Cancer research** nel campo titolo, e qui si evidenzia subito la caratteristica dell'**autocompletamento**, che è costante in Pubmed per cui è bene tenere sempre un occhio a quello che si digita perché di solito non è necessario completare la digitazione e ci vengono anche evitati eventuali errori.

Inserisco l'anno 2010 e qualche parola del titolo. Dovrebbe essere sufficiente, ed ecco il nostro articolo:

**Bianchini G**, Iwamoto T, Qi Y, Coutant C, Shiang CY, Wang B, Santarpia L, Valero V, Hortobagyi GN, Symmans WF, Gianni L, Puztai L. **Cancer Res.** **2010** Nov 1;70(21):8852-62. doi: 10.1158/0008-5472.CAN-10-1039. Epub 2010 Oct 19.

Posso tornare indietro e ripetere la ricerca in campi diversi, ad es. numero della pagina e autore: potrebbe essere sufficiente, ed ecco il nostro articolo.

## 2<sup>a</sup> prova

**Bianchini G**, Iwamoto T, Qi Y, Coutant C, Shiang CY, Wang B, Santarpia L, Valero V, Hortobagyi GN, Symmans WF, Gianni L, Puztai L. **Cancer Res.** 2010 Nov 1;70(21):**8852**-62. doi: 10.1158/0008-5472.CAN-10-1039. Epub 2010 Oct 19.

Il titolo della rivista può essere inserito in forma completa o abbreviata ... se digito il nome dell'autore mi compare un elenco che chiaramente non è in ordine alfabetico, ma di frequenza di ciascun autore all'interno della banca dati. Inserisco l'anno e tanto basta. Avrei trovato la citazione anche semplicemente con autore, anno volume e fascicolo o anche autore e numero di pagina.

Nel cercare il nome di un autore è più sicuro cercarlo con l'iniziale del nome anche se lo vedo con il nome intero: la ricerca con l'iniziale reperisce tutto, la ricerca solo con il nome intero reperisce solo le citazioni in cui l'autore è citato con il nome intero. E' chiaro che c'è sempre un rischio relativo alla disambiguazione rispetto ad autori con stesso cognome e iniziale, ma questo va risolto guardando altri aspetti, p.es l'istituto, l'area di interesse ...

Puliamo la maschera e facciamo dunque una ricerca per autore con il **citation matcher** per ottenere tutte le citazioni in cui è presente,

Provo prima **Zanotti Renzo** e poi **Zanotti R**

I risultati vengono visualizzati in Pubmed in un formato breve, il formato **summary**, per cambiare formato andare in alto a sinistra su **Display settings**.

[N.B. Il menù Display settings non esiste più: le opzioni che comprendeva sono ora tutte visibili sopra l'elenco dei risultati]

Ci sono diverse opzioni di visualizzazione: il formato **Abstract** è quello che ci permette di vedere il record completo. Si può scegliere il numero di records per pagina (di default 20) e anche l'ordinamento. Di default si visualizzano i record in ordine di pubblicazione dal più recente al meno recente, di recente introduzione è la visualizzazione per **ordine di rilevanza**. Introdotto nell'ottobre 2013, è basato su un algoritmo che attribuisce il "peso" delle citazioni in relazione al numero e al campo in cui sono presenti i termini di ricerca; anche la data di pubblicazione contribuisce al "peso" delle citazioni favorendo le citazioni più recenti.

## **Descrizione di una citazione bibliografica**

I records di Pubmed sono citazioni bibliografiche. Cos'è una citazione bibliografica? E' un insieme di informazioni che riportano dei dati essenziali relativi a un documento. I documenti di riferimento in Pubmed sono per la maggior parte articoli di riviste scientifiche pubblicati nelle riviste indicizzate in Pubmed. Ogni citazione bibliografica è formata da campi che contengono informazioni che ci dovrebbero aiutare a capire se tali articoli siano di nostro interesse e intendiamo perciò cercarne il testo completo per una lettura integrale.

Vediamo i campi dei records:

La prima riga contiene le coordinate dell'articolo:

- **Titolo abbreviato della rivista** (passandoci sopra il mouse si vede il titolo completo), anno, volume, fascicolo, pagina, eventualmente con la data di pubblicazione online; **N.B.** può essere presente l'indicazione tra parentesi

quadre **[Epub ahead of print]** cioè dice che l'articolo è stato pubblicato online e i dati bibliografici riportati **potrebbero non essere definitivi**.

- **Titolo dell'articolo [N.B. Se l'articolo originale non è in inglese ...]**

Se l'articolo non è in inglese il titolo è tra parentesi quadre mentre sotto è indicata la lingua dell'articolo originale, es. scriviamo NURSES DIAGNOSES = una delle prime citazioni ha il titolo tra parentesi quadre, se apriamo il titolo vediamo che sotto è specificata la lingua dell'originale che è la lingua giapponese.

- Torniamo alla nostra ricerca **Zanotti R[author]**
- Autori
- Aprendo **Author information** si aprono le informazioni sugli autori. Fino al 2013 le citazioni a controllo ultimato contenevano solo l'affiliazione del primo autore. Dal 2014 si vedono quelle di tutti gli autori.
- **Abstract:** presente solo se l'autore lo ha mandato in inglese
- **Keywords:** campo introdotto nel 2013 contenente parole chiave attribuite dagli autori (non hanno a che vedere con altre parole chiave, chiamate mesh che fanno parte di un vocabolario controllato e che vedremo in dettaglio)
- **PMID:** è un codice che identifica in modo univoco la citazione dentro a Pubmed, è sufficiente inserire questo codice nel box di ricerca per rintracciare una citazione, possiamo vederne un esempio copiando semplicemente il pmid di un articolo nel box di ricerca.

**Status:** accanto al PMID tra parentesi quadre è indicato lo status della citazione che può essere, p.es. **[PubMed - as supplied by publisher]**, questo è lo status degli articoli che, immediatamente alla pubblicazione online vengono mandati alla NLM che li inserisce in Pubmed, e ci dice che così come sono stati ricevuti vengono visualizzati prima di qualsiasi intervento di controllo.

altro status è [**PubMed - in process**], come dire “lavori in corso sull’articolo”; un altro ancora è [**Pubmed**], status di articoli che vengono ospitati nella banca dati, ma non vengono indicizzati (= non saranno loro attribuiti i mesh utili a ricercarli, non passeranno cioè allo status Pubmed –indexed for Medline).

Lo status che ci dice che la citazione è stata controllata formalmente e analizzata nei contenuti è [**PubMed - indexed for MEDLINE**]: tutte le citazioni che hanno questo status, il 90% di quelle presenti in Pubmed, hanno un ulteriore campo contenente altri elementi, tra i quali quello che ci interessa di più ai fini della ricerca bibliografica è il campo **Mesh** che va aperto: si tratta di un elenco di termini che fanno parte di un vocabolario controllato, il vocabolario Mesh, che indicano i concetti trattati nell’articolo. Alcuni sono contrassegnati dall’asterisco e segnalano i concetti su cui l’articolo si focalizza maggiormente.

Altro status che qui non troviamo è [**PubMed - OLDMEDLINE**]. Si tratta di circa 2 milioni di articoli pubblicati tra il 1946 e il 1965 indicizzati con criteri diversi da quelli attuali, cioè i Mesh usati non corrispondono al vocabolario Mesh attualmente in uso.

Non tutti i record passano alla fine allo status [**PubMed - indexed for MEDLINE**]

## **A PROPOSITO DI MEDLINE E PUBMED**

Si parla di Pubmed e Medline: qual è la differenza.

[SLIDE]

- **Pubmed** indica sia la piattaforma di visualizzazione disponibile gratuitamente dal 1996 che i contenuti dell’intero archivio
- **Medline** si riferisce al sottoinsieme di Pubmed costituito dalle citazioni indicizzate aventi status “*Indexed for Medline*” caratterizzate dalla presenza di

parole chiave chiamate Mesh appartenenti a un vocabolario controllato e da altri elementi che ne definiscono il contenuto (i record Medline costituiscono il 90% dell'archivio)

Continuando con la descrizione dei campi, abbiamo:

- **Related citations:** link a citazioni affini alla citazione da cui si parte. [da giugno 2015 rinominate **Similar articles**]
- **Altri links:** nel formato abstract può essere presente il link all'editore che permette di acquistare l'articolo, può essere presente il link a una versione free dell'articolo o può esserci il link all'eventuale open url resolver dell'istituzione. **Nel caso dell'università di Padova, in rete di ateneo (o da collegamento via proxy) ed entrando in Pubmed da un sito istituzionale è visibile l'icona AIREGO che permette di visualizzare i servizi presenti sulla citazione. Parlo di servizi perché non c'è soltanto il link a un eventuale full text dell'articolo, benché questo sia chiaramente il servizio più importante.**

Dal 2005 è cominciata la mappatura dei record precedenti al 1965 (oldmedline) secondo i criteri attuali. Lo status OLDMEDLINE resta solo fintanto che i mesh non sono stati tutti cambiati e mappati secondo i criteri correnti.

- **SUPPLEMENTARY CONCEPTS:** si tratta soprattutto di farmaci e sostanze chimiche indicati dagli indicizzatori perché trattati nell'articolo, sono descritti nel vocabolario Mesh e hanno il proprio riferimento a uno o più termini mesh senza avere una pagina articolata e dettagliata come i mesh.

Segue descrizione servizi accessibili con AIREGO (link al full text, link al catalogo del Sistema Bibliotecario Padovano in caso non ci sia la versione online, altri servizi, es. possibilità di esportare la citazione in Refworks, applicazione in uso

nell'università di Padova per gestire le bibliografie, e linka motori di ricerca affidabili (es. Scirus e Google Scholar) dove possiamo riproporre la nostra ricerca.