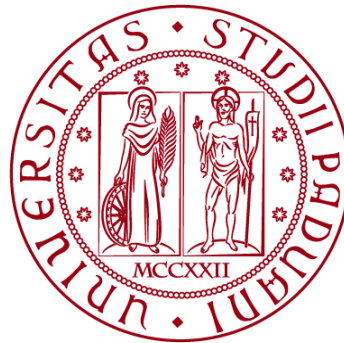


UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI BIOLOGIA

Corso di Laurea in Biotecnologie



ELABORATO DI LAUREA

**L'INTELLIGENZA ARTIFICIALE COME
STRUMENTO IN CAMPO
FARMACEUTICO**

**Tutor: Prof.ssa Adriana Chilin
Dipartimento di Scienze del Farmaco**

**Laureando: Tomas Florian
Matricola n. 2042940**

ANNO ACCADEMICO 2023/2024

INDICE

ABSTRACT.....	1
RICERCA E SVILUPPO IN CAMPO FARMACEUTICO.....	2
Di cosa si tratta.....	2
Limiti e svantaggi della R&S.....	2
L'INTELLIGENZA ARTIFICIALE.....	3
Che cos'è?.....	3
Tecniche e modelli di apprendimento.....	4
Applicazioni.....	6
IA IN CAMPO FARMACEUTICO.....	7
UN CASO COMPLICATO: LA FIBROSI POLMONARE IDIOPATICA.....	9
Di cosa si tratta.....	9
Cause e fattori di rischio.....	9
Terapie standard e limitazioni.....	11
L'INNOVAZIONE DI INSILICO MEDICINE: INS018_055.....	11
Panoramica sull'azienda.....	11
Tecnologie utilizzate.....	11
Alla ricerca di una cura per l'IPF.....	13
INS018_055.....	14
CONCLUSIONI.....	16
BIBLIOGRAFIA.....	18

ABSTRACT

La ricerca e lo sviluppo in campo farmaceutico sono processi necessari per portare sul mercato un nuovo farmaco. Tuttavia, questi processi sono molto dispendiosi sia in termini economici che di tempo impiegato e, nella maggior parte dei casi, le nuove molecole create rischiano di fallire durante la sperimentazione clinica.

Negli ultimi anni la potenza dell'Intelligenza Artificiale è aumentata. In particolare, nel campo farmaceutico si è iniziato ad utilizzare l'IA per identificare i possibili target farmaceutici e per ideare nuove molecole, impiegando tempistiche e risorse notevolmente inferiori rispetto all'approccio classico.

Una patologia complicata da trattare per colpa della sua eziologia è la fibrosi polmonare idiopatica (IPF), in quanto non sono stati ancora compresi i meccanismi molecolari che ne sono la causa.

Per trovare una soluzione a questa patologia, la Insilico Medicine ha impiegato i propri algoritmi di IA per identificarne il target e per sviluppare una nuova molecola (INS018_055). Le future prospettive su di essa sono promettenti, in quanto sono attualmente in corso test clinici di fase II.

Nonostante questi primi successi, l'Intelligenza Artificiale deve ancora essere validata come strumento in grado di sostituire gli approcci tradizionali in campo farmaceutico, ma i suoi potenziali ora sono tangibili e non più fantascienza.

RICERCA E SVILUPPO IN CAMPO FARMACEUTICO

DI COSA SI TRATTA

L'attività di Ricerca e Sviluppo (R&S) in campo farmaceutico è quell'insieme di processi che hanno come obiettivo principale l'introduzione nella pratica clinica e nel mercato di nuovi farmaci.

Essa è principalmente composta da cinque fasi: scoperta e validazione del target, individuazione e miglioramento della molecola affine al target, sviluppo della formulazione per studi preclinici, studi preclinici ed infine studi clinici.

Queste fasi possono essere quindi raccolte in tre macrocategorie: scoperta del farmaco, sviluppo del farmaco per studi preclinici e sviluppo del farmaco per un utilizzo clinico. Per poter iniziare gli studi clinici, cioè che coinvolgono direttamente individui umani, è assolutamente necessario che il farmaco abbia soddisfatto determinati requisiti di efficacia e sicurezza nelle fasi precedenti.

Gli studi clinici sono a loro volta suddivisi in quattro fasi. Nella fase I si studia la sicurezza del farmaco, nella fase II si valuta la sua efficacia, nella fase III si amplia lo studio a più centri di ricerca per poter confermare a livello statistico i risultati delle prime due fasi e per poterne affinare dosaggio e formulazione ed, infine, nella fase IV il farmaco, finalmente in commercio, è sottoposto ad un monitoraggio per confermare o negare ulteriormente gli aspetti sulla sua sicurezza.

Raramente un'azienda farmaceutica procede con lo sviluppo di una molecola oltre la seconda fase dei trial clinici se non si hanno tangibili evidenze che essa possa diventare un farmaco di successo, in quanto sarebbe un inutile dispendio di tempo e risorse economiche.

LIMITI E SVANTAGGI DELLA R&S

Per sviluppare un nuovo farmaco, le aziende farmaceutiche necessitano di investire ingenti risorse economiche. È stato calcolato che nel 2022 l'importo medio utilizzato dalle maggiori aziende farmaceutiche per sviluppare un nuovo farmaco si è attestato tra i 2-3 miliardi di dollari. Questo valore è dovuto ad un progressivo aumento delle difficoltà nello scoprire nuovi target terapeutici e anche dall'obiettivo di offrire terapie sempre più specifiche per ogni paziente. Per rendere l'idea sulla significatività dell'aumento di questa cifra, circa un decennio prima, in particolare nel 2010, la stima per la R&S di un nuovo farmaco era di circa 1 miliardo di dollari, quasi la metà di quanto stimato nel 2022.

Dell'intero budget richiesto, circa un terzo viene speso solamente nella fase di ricerca di base, in cui l'obiettivo è quello di comprendere il meccanismo biologico e/o molecolare che determina una determinata patologia. In questa fase i ricercatori

molto spesso devono procedere tramite una metodologia trial-and-error, che è intrinsecamente dispendiosa in termini di tempo e di denaro.

L'aspetto economico è solo una faccia della medaglia, in quanto è da considerare anche il tempo impiegato per lo sviluppo di un nuovo farmaco. Mediamente vengono impiegati tra i 10 e i 15 anni, una tempistica particolarmente elevata che, oltre a ritardare il ritorno economico da parte dell'azienda farmaceutica, preclude la possibilità di curare potenziali nuovi pazienti, soprattutto durante il periodo precedente all'introduzione della molecola negli studi clinici.

Per queste ragioni i ricercatori negli ultimi anni hanno iniziato a testare metodi innovativi di R&S che possano ridurre notevolmente tempistiche e costi, cosicché i vantaggi possano essere percepiti sia dalle aziende farmaceutiche che dagli individui malati. L'alternativa principale alle metodologie classiche è l'utilizzo dell'Intelligenza Artificiale, tecnologia applicabile in diverse fasi della R&S.

L'INTELLIGENZA ARTIFICIALE

CHE COS'È?

L'Intelligenza Artificiale (IA) è una tecnologia che consiste nell'applicazione di algoritmi informatici da parte di macchinari e computer per poter assolvere azioni che, tradizionalmente, necessitano dello sforzo mentale (anche detto "intelligenza") di un essere umano.

Nonostante sia un argomento di discussione mediatica degli ultimi anni, l'IA come concetto di base fu ideata negli anni '50, grazie all'ingegno di scienziati come Alan Turing (che ideò la prima "macchina intelligente", capace di simulare le risposte di una persona a determinate domande) e Arthur Samuel (che programmò un computer in grado di giocare a scacchi).

Negli anni a seguire questa tecnologia fu oggetto di studi e sperimentazioni che le permisero di "maturare" fino a raggiungere un periodo di massimo interesse da parte di un grande numero di aziende. Proprio in questo periodo, dal 1980 al 1987, ad esempio, vennero utilizzati a livello industriale chatbot e programmi capaci di selezionare le componenti necessarie per assemblare un computer in base alle richieste del cliente. Un altro evento significativo di questi anni fu lo sviluppo del primo veicolo a guida autonoma, realizzato da Ernst Dickmann in collaborazione con un team dell'Università di Monaco, in grado di raggiungere una velocità massima di 96 km/h su strada chiusa al traffico.

Tuttavia, a questo periodo di grande interesse per l'IA seguirono sei anni detti "AI Winter" in cui il suo sviluppo, metaforicamente, si congelò. La causa di questo

evento fu l'ampia sproporzione tra costi d'investimento e ritorni economici, principalmente dovuta all'assenza di hardware più performanti in grado di permettere un più ampio e migliore utilizzo delle tecniche di Intelligenza Artificiale.

Con la fine degli anni '90 e l'inizio del nuovo millennio, si è assistito a un progressivo avanzamento tecnologico nell'hardware, ad esempio nelle GPU (Graphic Processing Units) e nella RAM (Random Access Memory), e l'Intelligenza Artificiale cominciò nuovamente ad essere vista come una tecnica promettente. Alcune tappe importanti del suo progresso in questa nuova fase furono: la sconfitta del pluricampione mondiale di scacchi Garry Kasparov contro il computer Deep Blue, nel 1997; nel 2002 fu rilasciato il primo modello di Roomba, un'aspirapolvere autonomo; nel 2004 la Nasa fece atterrare su Marte due rover (dei veicoli per l'esplorazione planetaria) con funzione di navigazione semi-autonoma; nel 2011 Apple rilascia il primo assistente virtuale utilizzabile dal pubblico: Siri.

Gli ultimi esempi citati servono ad evidenziare come l'IA, con il passare degli anni, si sia evoluta non solo verso un utilizzo prettamente lavorativo ma anche per la fruizione da parte dell'intera popolazione.

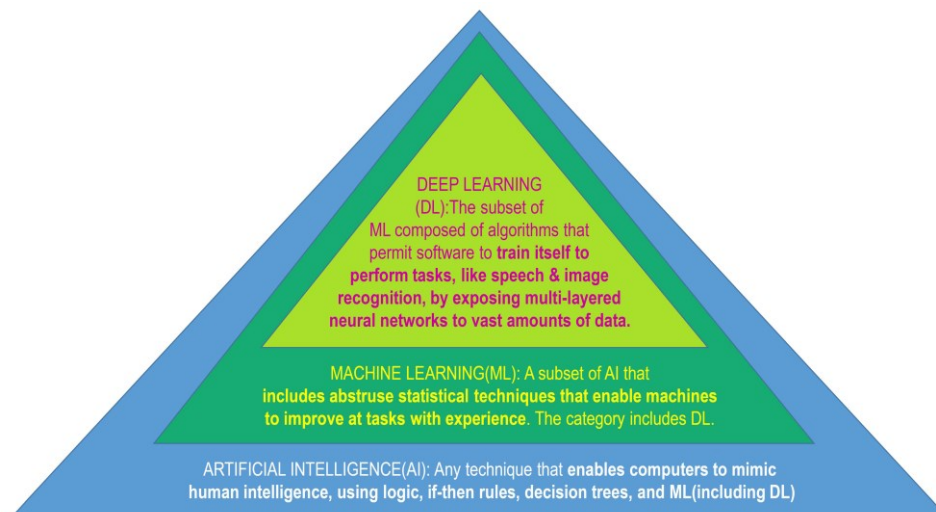
Con il nuovo millennio, oltre ai progressi hardware, emerse anche il fenomeno dei "Big Data", che rese possibili l'acquisizione e l'archiviazione di enormi quantità di dati digitali, potenzialmente accessibili da qualsiasi parte del mondo grazie a Internet.

L'ultima fase dell'intelligenza artificiale, iniziata approssimativamente nel 2012, è caratterizzata dall'integrazione dell'IA nella vita quotidiana delle persone e dallo sviluppo avanzato delle tecniche di deep learning.

TECNICHE E MODELLI DI APPRENDIMENTO

L'Intelligenza Artificiale non è una capacità intrinseca di un computer o di un macchinario, ma è il risultato di una combinazione di elementi che permettono ad un sistema informatico di apprendere da un insieme di dati per poi applicare le conoscenze acquisite a dei casi concreti.

I termini "Intelligenza Artificiale" non indicano una singola tecnica, bensì la più ampia macrocategoria di questo tipo di tecnologia. Esistono quindi numerose tecniche di IA, suddivisibili in molteplici categorie, specifiche per gli algoritmi utilizzati e le tipologie di problemi da risolvere. Alcuni esempi sono le tecniche di elaborazione del linguaggio, quelle di visione artificiale, i sistemi di raccomandazione e quelli di apprendimento. Questa ultima categoria è la più utilizzata in ambito scientifico e di ricerca e si suddivide in due tipologie principali: apprendimento automatico e apprendimento profondo, più noti con i termini inglesi *Machine Learning* e *Deep Learning*.



Artificial Intelligence and Machine Learning Technology Driven Modern Drug Discovery and Development.
Int. J. Mol. Sci. 2023, 24, 2026. <https://doi.org/10.3390/ijms24032026>

La più longeva e semplice tra le due è il Machine Learning (ML), una tecnica di IA che sfrutta uno o più algoritmi per analizzare e apprendere informazioni dai dati presenti all'interno di un database, per poi formulare previsioni ponderate su quanto appreso.

Il Deep Learning (DL), una sottocategoria di ML, è una tecnica più potente e più complessa che permette di analizzare database di dimensioni maggiori e di creare connessioni logiche basate sulla similarità di pattern ricorrenti, per poi dedurre la regola che esiste potenzialmente alla base di ogni singola informazione. Questo lavoro segue il modello delle “reti neurali artificiali” (ANN, dall'inglese Artificial Neural Network) che mimano il sistema del ragionamento umano in modo ordinato e schematico, qualunque siano le informazioni da analizzare.

Le ANN sono composte da un insieme di nodi, ciascuno dei quali riceve uno o più input, i quali vengono poi trasformati in output. Questi nodi possono operare individualmente o essere interconnessi, a seconda dell'algoritmo utilizzato per l'elaborazione dei dati.

Per addestrare un algoritmo, sia che faccia parte di un sistema di ML che di DL, si possono seguire quattro diversi modelli: supervisionato, non supervisionato, parzialmente supervisionato e di rinforzo.

L'addestramento supervisionato avviene grazie all'utilizzo di database contenenti dati e variabili sia indipendenti che dipendenti. Un esempio di variabili indipendenti nel campo medico-farmaceutico possono essere le proprietà fisico-chimiche e strutturali di una molecola, mentre le variabili dipendenti possono essere l'attività biologica o altre proprietà della molecola d'interesse. Avendo a disposizione entrambe le variabili, l'algoritmo potrà generare modelli di regressione in grado di correlare i due insiemi di dati.

L'addestramento non supervisionato, invece, avviene in mancanza delle variabili dipendenti e, quindi, l'algoritmo dovrà essere in grado di produrre ipotesi su di esse. Questo sistema è fondamentale negli strumenti di IA utilizzati in campo farmaceutico, in quanto le informazioni associate ad alcune patologie, come i target e i pathway biologici, spesso sono sconosciuti.

L'addestramento parzialmente supervisionato è utile quando sono disponibili solamente un numero ridotto di dati da poter fornire all'algoritmo. Questo meccanismo è una via intermedia tra i due metodi sopracitati, in quanto la metodica prevalente sarà quella non supervisionata che sarà però parzialmente guidata dai pochi dati già esistenti.

L'apprendimento di rinforzo può essere supervisionato o meno. Nel secondo caso, sarà compito dell'algoritmo dover assegnare un feedback negativo o positivo alle scelte da esso stesso precedentemente effettuate, in modo da apprendere un metodo di decisione basato su percorsi che possano condurre all'ottenimento dei risultati più attendibili.

APPLICAZIONI

Dopo aver addestrato e validato un algoritmo, è possibile sfruttare il sistema di IA per esaudire le richieste dell'operatore. Al giorno d'oggi esistono poche mansioni che non siano ottimizzabili o addirittura automatizzabili grazie a questa tecnologia.

Nel campo finanziario e azionario, ad esempio, le analisi di mercato possono essere facilitate da algoritmi in grado di prevedere l'andamento futuro delle azioni, basandosi su quanto successo nel passato. Il servizio di assistenza clienti di una qualsiasi azienda, come una produttrice di elettrodomestici o un operatore di telefonia mobile, può essere completamente automatizzato attraverso l'utilizzo di bot capaci di comprendere le esigenze del cliente e di indirizzarlo verso la soluzione più adatta; il coinvolgimento di un operatore umano sarà necessario solo in caso di richieste complesse o particolarmente specifiche.

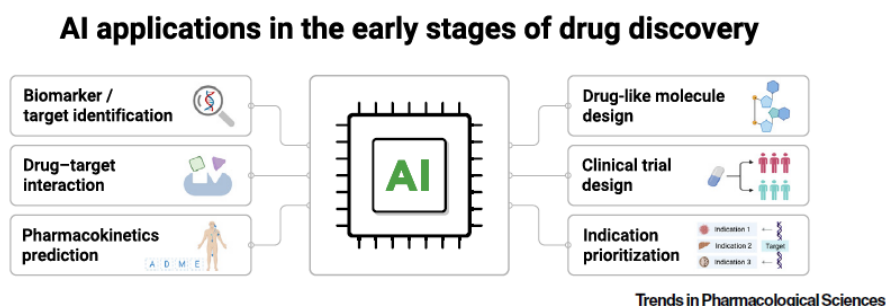
Questi sono solo alcuni esempi, ma ciò che l'utilizzo dell'Intelligenza Artificiale permette in tutti quanti i casi è il risparmio di tempo e lavoro umano richiesto, rendendo più efficienti anche quelle mansioni che precedentemente richiedevano uno smisurato impegno. Nel mondo del lavoro, tutto ciò si riflette in un minor dispendio di risorse e, quindi, in maggiori ricavi da parte delle aziende.

Esistono, tuttavia, situazioni in cui l'adozione dell'IA non è motivata esclusivamente da un incremento dei profitti. Basti pensare, ad esempio, al campo medico, dove questa tecnologia potrebbe accelerare e rendere più accurata una diagnosi, oppure a quello farmaceutico, in cui sarebbe possibile ridurre drasticamente i tempi necessari per sviluppare terapie innovative. In entrambi i casi, l'IA sarebbe in grado di offrire ai pazienti migliori prospettive di guarigione.

IA IN CAMPO FARMACEUTICO

Nel settore farmaceutico, l'Intelligenza Artificiale trova applicazione soprattutto nei processi di ricerca e sviluppo, che, come accennato nei primi paragrafi, si sono dimostrati particolarmente onerosi sia in termini di tempo che di risorse economiche.

I progressi degli ultimi decenni nel campo dell'IA e, in particolare, del Deep Learning hanno portato alla creazione di nuovi algoritmi e programmi capaci di svolgere le numerose attività richieste dalla R&S.



AI-powered therapeutic target discovery. doi.org/10.1016/j.tips.2023.06.010

Gran parte delle patologie che affliggono l'essere umano sono dovute a mutazioni o disfunzioni di proteine, il che rende queste ultime, nella maggior parte dei casi, i target terapeutici per eccellenza.

Sebbene ormai sia nota la sequenza primaria di moltissime proteine, quella secondaria e quella terziaria sono ancora per lo più ignote. Questo è dovuto alla proprietà intrinseca delle proteine di essere entità dinamiche e, quindi, difficilmente caratterizzabili in modo univoco. Per ovviare a questo limite esistono approcci di DL sempre più affidabili in grado di determinare le plausibili strutture tridimensionali delle proteine a partire dalla semplice struttura primaria.

In aggiunta a questo limite conoscitivo, si aggiunge anche il fatto che le proteine non agiscono come molecole isolate, ma mantengono la loro stabilità ed esercitano le loro funzioni attraverso interazioni con altre proteine o molecole. Anche in questa circostanza, l'Intelligenza Artificiale si dimostra particolarmente efficace nel prevedere queste associazioni, valutandone l'efficacia, il successo dell'interazione e la possibilità che la proteina d'interesse possa essere bersaglio di un farmaco già esistente.

Quest'ultimo approccio permette la riproposizione di un farmaco per un nuovo utilizzo e, quindi, di accelerare l'avvio degli studi clinici per verificarne l'efficacia nel nuovo contesto di utilizzo, riducendo significativamente i costi che altrimenti sarebbero stati necessari per lo sviluppo *de novo* di una nuova molecola e tutti gli studi necessari a permetterne l'utilizzo clinico.

Gli studi sulle associazioni farmaco-proteina, oltre a riproporre farmaci per nuovi utilizzi, potrebbero anche individuare eventuali capacità di una molecola di legare diversi recettori e quindi indurre possibili reazioni avverse off-target. Queste scoperte consentono ai ricercatori di focalizzare il lavoro sull'ottimizzazione del farmaco, migliorandone la sicurezza e la selettività verso il target desiderato. Anche questo processo può essere supportato dall'Intelligenza Artificiale.

Se invece il sistema non individua farmaci da riproporre, si può ricorrere ad altri algoritmi capaci di generare molecole *de novo*, considerando innanzitutto i legami che dovrebbe effettuare con il target d'interesse e quindi anche la conformazione spaziale. In seguito, bisogna considerare anche altre proprietà caratterizzanti il profilo farmacocinetico del farmaco, che implicano quindi i processi di assorbimento, distribuzione, metabolismo (o per meglio dire, biotrasformazione), escrezione e tossicità, senza però dimenticare che la nuova molecola deve poter essere sintetizzata. Conoscere questi aspetti prima dell'effettiva sintesi della molecola può limitare, o addirittura evitare, a priori il fallimento delle successive sperimentazioni *in vivo*.

La sintesi di una nuova molecola è un processo complesso che richiede un'attenta pianificazione delle reazioni chimiche necessarie per trasformare dei reagenti nella molecola desiderata. Spesso, il risultato non è direttamente la molecola di interesse, ma un insieme di precursori più semplici e facilmente ottenibili. Questo approccio, tuttavia, non parte dall'analisi dei reagenti, ma procede al contrario: si scompone la molecola target nei suoi precursori, identificando successivamente le reazioni chimiche necessarie per ottenerli. Per questo motivo, questa metodologia è definita retrosintesi, più che una vera e propria sintesi diretta.

Per facilitare e rendere più rapida questa analisi, esistono algoritmi di IA, come il MCTS (Monte Carlo Tree Search) o il più recente 3N-MCTS, che sono in grado di prevedere l'efficacia delle ipotetiche reazioni per poi selezionare solamente quei percorsi che sono risultati ottimali e con una limitata produzione di molecole indesiderate.

Dopo aver deciso il set di reazioni che devono occorrere, dei sistemi robotizzati e governati da precisi e validati algoritmi (un esempio è Chemplier) potranno eseguire la sintesi delle nuove molecole senza richiedere l'intervento umano. Oltre a non richiedere lavoro manuale e a ridurre le tempistiche, questi sistemi offrono il principale vantaggio di eseguire procedure in modo estremamente preciso, sicuro e in modo replicabile.

Oltre a semplificare lo sviluppo di nuovi composti, l'IA può anche essere utilizzata per decidere gli schemi per condurre i trial clinici, selezionando gli individui più adatti ad essere sottoposti allo studio, evitando in questo modo di ottenere troppi dati ripetitivi o di escludere dai risultati una parte di popolazione diversa, ad

esempio, per stili di vita, condizioni fisiche, etnia o altre caratteristiche che difficilmente verrebbero considerate.

Tutti questi concetti ad un osservatore non esperto potrebbero sembrare molto astratti. Per rendere tangibile a tutti l'utilità e la potenza dell'IA nel processo di R&S di farmaci, specialmente per patologie gravi e poco conosciute, può essere utile esaminare il caso della fibrosi polmonare idiopatica e le strategie adottate dall'industria farmaceutica per affrontarla.

UN CASO COMPLICATO: LA FIBROSI POLMONARE IDIOPATICA

DI COSA SI TRATTA

La fibrosi polmonare idiopatica (IPF) è una grave patologia che affligge i polmoni, distruggendone il tessuto e la struttura. In particolare, si assiste ad una deposizione incontrollata di collagene e altri componenti della matrice extra-cellulare a livello di alveoli e bronchioli, strutture costituenti la parte funzionale del polmone: il parenchima.

L'eccessivo deposito di collagene in queste strutture causa nei pazienti sintomi tra cui tosse secca, dispnea e affaticamento respiratorio, in quanto gli alveoli e i bronchioli, essendo ostruiti, non permettono un sufficiente scambio e flusso di gas.

Questa patologia si verifica principalmente in individui con più di sessanta anni che, nella maggior parte dei casi, vengono diagnosticati in uno stadio già avanzato della malattia, con scarsissime probabilità di guarigione.

CAUSE E FATTORI DI RISCHIO

Le cause che portano allo sviluppo di IPF non sono ancora state completamente scoperte e dipendono da una moltitudine di fattori, che possono essere sia ambientali che genetici, oltre ad una risposta variabile in ogni singolo paziente.

Per quanto riguarda la predisposizione genetica possono essere distinte due situazioni: quella più frequente nella popolazione è la presenza di SNPs (polimorfismi di singolo nucleotide), con una frequenza allelica maggiore dell'1%, e quella meno frequente è determinata da varianti rare con frequenza minore dell'1%.

Queste mutazioni sono state scoperte negli ultimi decenni grazie all'avvento delle nuove generazioni delle tecnologie di sequenziamento (NGS) che hanno reso più semplice ed economico il sequenziamento dei genomi. Per individuare quali fossero

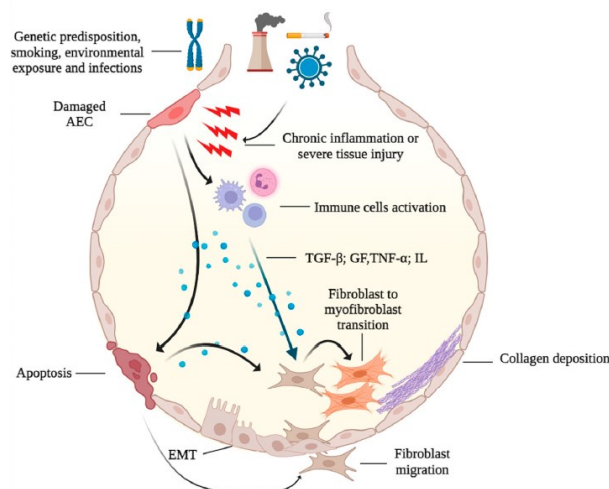
le possibili mutazioni, gli scienziati si sono concentrati sull'analizzare i pathway del mantenimento dei telomeri e quello del metabolismo del surfattante polmonare.

Per il mantenimento dei telomeri è fondamentale l'attività delle telomerasi, classe di enzimi che aggiunge sequenze ripetute alla fine dei cromosomi in prevenzione del naturale processo di accorciamento. Le subunità catalitiche richieste sono la TERT (trascrittasi inversa delle telomerasi) e il TERC (componente di RNA delle telomerasi) ed è stato osservato che negli individui affetti da IPF, l'attività di queste subunità è alterata.

Il surfattante polmonare è una sostanza formata dall'insieme di lipidi e proteine che viene prodotta dai pneumociti di tipo II. Questa sostanza ha il ruolo principale di ridurre la tensione superficiale degli alveoli, che sono le strutture deputate allo scambio gassoso con la circolazione ematica. Se il surfattante non fosse presente negli alveoli, essi collasserebbero. Tuttavia, alcuni problemi possono insorgere anche quando la sua produzione avviene in maniera eccessiva e sregolata. Questo accade, ad esempio, quando un individuo è soggetto a mutazioni sul gene MUC5B, responsabile della produzione di un componente della mucosa (la mucina 5B).

La probabilità di sviluppare la IPF può dipendere, oltre che dalla predisposizione genetica, anche dallo stile di vita dell'individuo. In particolare, le principali azioni che possono portare all'insorgere di questa patologia sono il lavorare in ambienti in cui vengono prodotti polveri sottili, fibre volatili e materiale particolato (sostanza prodotta anche dai fumi di scarico dei veicoli a combustione) e soprattutto il fumare sigarette. È stato dimostrato che i fumatori abituali sono soggetti il 60% in più ad essere affetti da IPF rispetto ai non fumatori.

Oltre ai danni ai tessuti e alle mucose respiratorie, queste sostanze possono anche indurre modificazioni epigenetiche nel DNA e nell'RNA. L'insieme dei danni fisici e genetici sono rilevanti nello sviluppo di gravi forme di IPF.



Highlights on Future Treatments of IPF: Clues and Pitfalls. Int. J. Mol. Sci. 2024, 25, 8392.
<https://doi.org/10.3390/ijms25158392>

TERAPIE STANDARD e LIMITAZIONI

Attualmente esistono due farmaci utilizzati come gold standard nel trattamento di IPF: Pirfenidone e Nintedanib. Essi, tuttavia, non sono in grado di curare o prevenire l'insorgere della malattia, ma riescono solamente a rallentare il progresso della malattia.

Il Pirfenidone è in grado di ridurre l'espressione e l'attivazione del fattore di crescita tumorale TGF- β , con la conseguente riduzione della sintesi di collagene e minore attivazione di fibroblasti, oltre ad un'attività antiinfiammatoria.

Il Nintedanib, invece, è un inibitore delle tirosin-chinasi in grado di legare e inibire recettori con un ruolo rilevante nell'espressione dell'IPF, come il VEGFR, il FGFR e il PDGFR.

Il motivo per cui non sono ancora in commercio farmaci più mirati è che non è ancora stato chiarito completamente il completo pathway molecolare e genetico.

Al momento sono sotto sperimentazione clinica diversi nuovi potenziali farmaci, la maggior parte dei quali è però ancora in fase I.

L'INNOVAZIONE DI INSILICO MEDICINE: INS018_055

PANORAMICA SULL'AZIENDA

Insilico Medicine è un'azienda biotecnologica improntata in ricerca e sviluppo di nuovi farmaci. L'approccio che l'azienda utilizza, e che è stato in grado di farle ottenere rilevanti successi, è decisamente innovativo e si basa sull'utilizzo dell'Intelligenza Artificiale per ridurre al minimo le sperimentazioni precliniche sugli animali e per accelerare le tempistiche richieste per la R&S dei farmaci.

Fondata nel 2014 negli Stati Uniti dal ricercatore e imprenditore Alex Zhavoronkov, Insilico Medicine mantiene la propria sede principale negli Stati Uniti. L'azienda ha inoltre un'importante sede operativa a Hong Kong, che le consente di condurre studi clinici su una popolazione culturalmente diversa da quella occidentale. Questa presenza in Asia è inoltre strategica per espandere la portata della ricerca e adattare le terapie a diverse esigenze sanitarie globali.

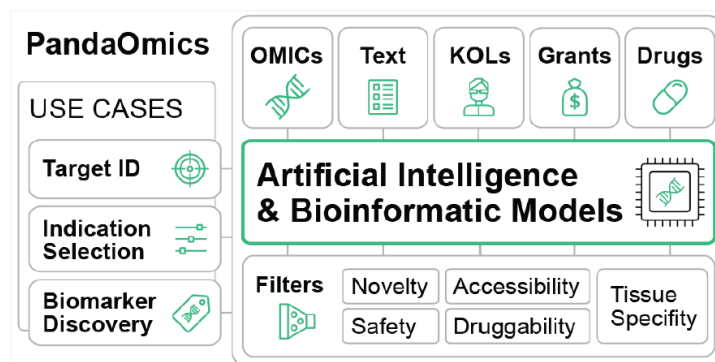
TECNOLOGIE UTILIZZATE

Prima di procedere con la spiegazione di come l'azienda è riuscita ad ottenere un farmaco potenzialmente utile nel trattamento dell'IPF, è importante definire quali

sono le due tecnologie che hanno permesso e che permetteranno sempre più di rendere più efficiente il processo di ricerca e sviluppo di nuovi farmaci per patologie orfane o di difficile comprensione. Le piattaforme in questione sono PandaOmics e Chemistry42, entrambe basate sull'utilizzo di sofisticati algoritmi di IA.

PandaOmics è una piattaforma informatica che permette di individuare nuovi target terapeutici basandosi sull'analisi di diversi database contenenti dati biologici, omici, clinici, chimici, pubblicazioni scientifiche e tutto ciò che può essere attinente al target che si vuole scoprire. Questa tecnologia è particolarmente rilevante per patologie complesse in cui è fondamentale individuare il target specifico che ne è la causa per poter sviluppare terapie curative in aggiunta a quelle palliative già esistenti. L'utente di questa piattaforma può filtrare l'output della ricerca secondo diversi criteri del target, come la sicurezza, la tessuto-specificità, la classe proteica di appartenenza, i pathway biologici in cui è coinvolto e la possibilità di interagire con farmaci o anticorpi. PandaOmics, tuttavia, non è un predittore assoluto di target terapeutici, ma è in grado di fornire un range di ipotesi, permettendo ai ricercatori di proseguire con l'approccio che ritengono più opportuno.

Anche se sono state riportate importanti evidenze di efficacia predittiva di PandaOmics, esso non è particolarmente efficace nel caso di patologie indotte da cause esterne, come infezioni virali o fungine con meccanismi molecolari non ben definiti.

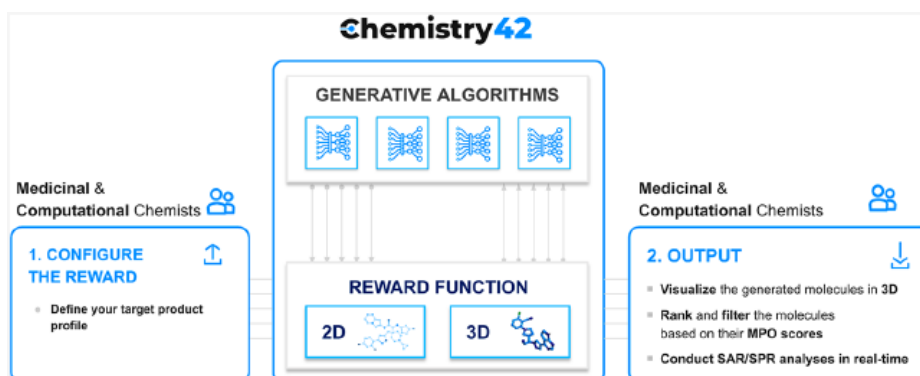


PandaOmics: An AI-Driven Platform for Therapeutic Target and Biomarker Discovery - Journal of Chemical Information and Modeling **2024** 64 (10), 3961-3969 DOI: 10.1021/acs.jcim.3c01619

Chemistry42 è invece una piattaforma che permette di progettare e generare *de novo* nuove molecole, basandosi su modelli di machine e deep learning in grado di valutare parametri chimici e biologici e di ottimizzare le proprietà di molecole già esistenti. Per adempiere la generazione di una nuova molecola, il sistema può seguire due approcci: LBDD o SBDD, rispettivamente *Ligand* e *Structure Based Drug Design*. Nel primo caso è richiesto il caricamento da parte dell'utente della struttura bidimensionale o tridimensionale del ligando, una stringa SMILE (un modo semplificato per rappresentare alfanumericamente una molecola) oppure è possibile disegnare manualmente la molecola nell'interfaccia di Chemistry42. Nel caso dell'approccio SBDD, invece, è richiesto l'inserimento della struttura della

proteina target. In entrambi i casi, l'utente può fornire particolari indicazioni sulle caratteristiche da farmacoforo della molecola da generare.

Come nel caso di PandaOmics, il ricercatore può scremare l'output tramite l'applicazione di filtri.



Chemistry42: An AI-Driven Platform for Molecular Design and Optimization - Journal of Chemical Information and Modeling **2023** 63 (3), 695-701 DOI: 10.1021/acs.jcim.2c01191

Sebbene PandaOmics e Chemistry42 siano due piattaforme separate, il loro vero potenziale si ottiene grazie al loro utilizzo combinato, in quanto PandaOmics può individuare i target di maggiore rilevanza ed in seguito Chemistry42 può progettare o ottimizzare composti specifici ed altamente affini per essi.

ALLA RICERCA DI UNA CURA PER L'IPF

Vista la scarsità di soluzioni terapeutiche per trattare gli individui affetti dalla fibrosi polmonare idiopatica, nell'agosto del 2019 i ricercatori della Insilico Medicine decisero di iniziare un processo di R&S utilizzando le piattaforme di intelligenza artificiale di loro proprietà, non partendo dal problema terapeutico in sé, ma bensì dall'individuazione del target specifico che ne è la causa.

Per iniziare il progetto di ricerca, i ricercatori iniziarono la piattaforma PandaOmics alla ricerca di un plausibile target, utilizzando database contenenti: dati multiomici prodotti dall'analisi di tessuti di pazienti affetti da IPF; informazioni sui trattamenti clinici effettuati; dati della letteratura scientifica. Grazie ad essi, PandaOmics è stato in grado di correlare i diversi tipi di dati e di stilare una lista di possibili target terapeutici, ordinati dal più al meno plausibile.

Affinché il target scelto dall'algoritmo possa essere oggetto di studio per una nuova molecola ad esso affine, è necessario essere il più statisticamente sicuri che la piattaforma sia in grado di effettuare le giuste connessioni ed ipotesi. Per questo motivo il metodo di ricerca utilizzato da Insilico Medicine è stato validato tramite un approccio "Time Machine", ovvero "allenando" la piattaforma tramite l'utilizzo di dati pubblicati entro una certa data per verificare se l'algoritmo è in grado di

ottenere come output lo stesso target che è stato scelto dall'industria farmaceutica dopo tale data.

I ricercatori decisero di concentrare la ricerca sui target coinvolti nella trasduzione dei segnali cellulari mediati da chinasi in una situazione di fibrosi. Il primo risultato individuato da PandaOmics fu TNIK, una chinasi appartenente alla famiglia delle serina-treonina chinasi. Questo risultato fu cruciale, poiché fino a quel momento per il trattamento dell'IPF erano stati sviluppati solo farmaci inibitori delle tirosin-chinasi, come il Nintedanib. Il ruolo delle serina-treonina chinasi, tuttavia, non era ancora stato approfondito né compreso appieno.

Successivi studi rivelarono che TNIK è coinvolto in processi chiave della fibrosi, come la differenziazione dei miofibroblasti, la migrazione delle cellule mesenchimali e il signaling dell'adesione focale. L'attivazione di questi meccanismi da parte di TNIK è stata dimostrata grazie ad un algoritmo di IA in grado di generare *de novo* i pathway molecolari che coinvolgono la chinasi.

Una volta individuato il target d'interesse, per individuare una molecola che lo potesse inibire i ricercatori utilizzarono i modelli preesistenti delle TNIK e la piattaforma Chemistry42. In particolare, decisero di generare una molecola in grado di legare con alta affinità il sito di legame per l'ATP, strategia molto utilizzata per inibire l'attività chinasi. Questa molecola necessitava di formare legami idrogeno con i residui amminoacidici della regione cerniera di TNIK e, per aumentare la selettività, anche con regioni allosteriche meno conservate adiacenti al sito attivo di legame.

La prima lista di molecole generate fu stilata selezionando solo quelle che rispettavano determinati requisiti, come novità, proprietà chimiche e la possibilità di essere sintetizzate. Le molecole generate dall'algoritmo furono in seguito sintetizzate e studiate in vitro; la maggior parte aveva affinità per il target a livelli nano-molari, ma tutte peccavano sulle quattro proprietà farmacocinetiche (ADME), rivelando nei microsomi umani e murini una elevata clearance e una bassa solubilità. Una volta individuate queste limitazioni si proseguì con una fase di ottimizzazione, il cui risultato fu il composto INS018_055. Per raggiungere questo risultato furono necessari solamente 18 mesi, un tempo record considerando che lo studio

INS018_055

Questa nuova molecola è risultata capace di formare legami idrogeno sia con la regione cerniera che con le aree adiacenti, contribuendo così alla stabilità del complesso molecola-target. Inoltre, INS018_055 può penetrare profondamente nella tasca del sito allosterico, contribuendo così ad aumentare l'affinità per il target che è caratterizzata da una costante di dissociazione di 4.32 nM, un valore molto

basso che indica la concentrazione di INS018_055 necessaria per occupare il 50% i siti di legame.

Per poi poter affermare che il composto potesse avere un ruolo determinante nel trattare la fibrosi, furono effettuati studi su fibroblasti polmonari con elevati livelli di TGF- β e, di conseguenza, una over espressione della proteina α -SMA, coinvolta nella differenziazione dei fibroblasti in miofibroblasti attivi e in grado di secernere componenti della matrice extracellulare. I primi studi in vitro dimostrarono che INS018_055 è in grado di ridurre i livelli di espressioni di TGF- β in relazione alla dose somministrata.

In seguito ne fu testata l'efficacia su fibroblasti provenienti da tre individui sani e da tre individui affetti da IPF. Nelle cellule provenienti dai malati si notò che il composto era in grado di inibire completamente e in modo dosaggio dipendente l'espressione di α -SMA; in quelle provenienti dagli individui sani si notò la stessa situazione, ma le concentrazioni di farmaco necessarie per indurre l'inibizione furono maggiori rispetto a quelle necessarie nei malati. Questa osservazione suggerì che INS018_055 avesse più efficacia negli individui affetti da IPF.

È stato inoltre osservato che nell'inibire la transizione delle cellule epiteliali in mesenchimali, in individui affetti da IPF il nuovo farmaco ha una concentrazione inibitoria molto inferiore rispetto a quella richiesta con l'uso di Nintedanib. In questo studio fu osservato che le cellule non venivano uccise, indice di bassa citotossicità della molecola.

Per avere ulteriori evidenze sull'efficacia anti-fibrotica di INS018_055 furono inoltre effettuati dei test in vivo su topi in cui venne indotta manualmente una disfunzione polmonare con l'utilizzo di bleomicina, un farmaco solitamente utilizzato con funzione antitumorale ma che, a dosaggi elevati e somministrazioni ripetute, può indurre come principale effetto indesiderato la fibrosi polmonare. Dopo due settimane di trattamento, i risultati furono ottimi dato che INS018_055 dimostrò una riduzione dosaggio dipendente degli effetti negativi sulla respirazione. Successivamente fu valutata anche l'attività antinfiammatoria.

Gli ottimi risultati ottenuti nella fase preclinica permisero quindi ad Insilico Medicine di proseguire con la sperimentazione clinica. Inizialmente venne condotta una fase 0, in cui dei partecipanti sani vennero sottoposti a somministrazioni di microdosi per valutare le proprietà farmacocinetiche della molecola nell'organismo umano.

In seguito venne avviata una sperimentazione in fase I in doppio cieco e, come controllo, contro placebo in individui sani volontari. L'obiettivo principale della fase I è stato quello di valutare la sicurezza e la tollerabilità di INS018_055, oltre che di riconfermarne i parametri farmacocinetiche. Inoltre, per verificare eventuali differenze in popolazioni culturalmente diverse, questi studi vennero effettuati sia in Nuova Zelanda che in Cina; da questo ne derivò che non ci fossero notevoli

differenze di comportamento del farmaco tra le due popolazioni, con buone proprietà farmacocinetiche, pochi effetti indesiderati (soprattutto non gravi) ed elevata sicurezza nell'assunzione.

Attualmente INS018_055 è in due studi clinici di fase II, sancendo il primato come primo farmaco disegnato interamente (a partire dalla scoperta del target molecolare) dall'IA ad entrare in questa fase degli studi clinici.

CONCLUSIONI

Nonostante i primi risultati degni di nota ottenuti nel campo farmaceutico grazie all'Intelligenza Artificiale, il suo potenziale utilizzo su scala industriale è ancora relativamente distante. Ciò che manca affinché la comunità scientifica possa affermare che questa tecnologia sia statisticamente affidabile ed efficace sono, appunto, un maggior numero di studi che abbiano introdotto nella pratica clinica nuovi farmaci.

Purtroppo, per raggiungere questo obiettivo sono necessari tempo, sperimentazioni e una raccolta sempre maggiore di dati omici e di strutture molecolari e proteiche, oltre che lo sviluppo di metodologie in grado di validare le tecnologie utilizzate.

Quest'ultimo punto è quello maggiormente attenzionato dalle agenzie regolatrici e di controllo dei composti farmaceutici, in quanto richiedono completa trasparenza sugli algoritmi, sui metodi di addestramento e sulle fonti dei dati utilizzati.

Questi limiti nello sviluppo dell'IA, tuttavia, sono uno stimolo aggiuntivo per le aziende a progredire con gli studi e velocizzare la possibilità di sfruttare l'IA per ottenere maggiori guadagni, temporali ed economici.

Grandi passi avanti sono stati fatti da Insilico Medicine con INS018_055, ma per poter trarre le conclusioni decisive sull'utilizzo di questo nuovo approccio di ricerca bisogna innanzitutto attendere la fine della sperimentazione clinica di fase II e, in caso positivo, anche quelle di fase III e IV.

Un ulteriore limite che potrebbe insorgere in un futuro in cui l'IA in questo campo venga accettata dalla comunità scientifica potrebbe essere il dover convincere anche le persone comuni ad accettare di sottoporsi a trattamenti con farmaci da essa generati. Infatti, gran parte delle persone, soprattutto nella fascia adulta ed anziana, tendono ad interpretare l'innovazione come qualcosa di non corretto e non necessariamente utile. Nel caso dell'IA, questo fenomeno potrebbe essere ancor più accentuato dalla "fobia" che i computer possano prendere il sopravvento sull'essere umano.

Per mitigare questa avversione, l'unica soluzione plausibile sarebbe quella di condurre campagne di informazione attraverso i maggiori canali mediatici, in modo da raggiungere più persone possibili e confutare le ideologie sbagliate e scientificamente infondate.

Detto questo, l'IA come strumento nella R&S deve ancora attendere del tempo per sostituire completamente l'approccio tradizionale, ma ormai le basi sono più che fondate e la strada per il suo successo comincia a mostrarne la fine o, per meglio dire, una nuova era dell'industria farmaceutica.

BIBLIOGRAFIA

1. “A small-molecule TNIK inhibitor targets fibrosis in preclinical and clinical models”.

Feng Ren, Alex Aliper, Jian Chen, Heng Zhao, Sujata Rao, Christoph Kuppe, Ivan V. Ozerov, Man Zhang, Klaus Witte, Chris Kruse, Vladimir Aladinskiy, Yan Ivanenkov, Daniil Polykovskiy, Yanyun Fu, Eugene Babin, Junwen Qiao, Xing Liang, Zhenzhen Mou, Hui Wang, Frank W. Pun, Pedro Torres-Ayuso, Alexander Veviorskiy, Dandan Song, Sang Liu, Bei Zhang, Vladimir Naumov, Xiaoqiang Ding, Andrey Kukhareenko, Evgeny Izumchenko & Alex Zhavoronkov

Published online: 08 March 2024. doi.org/10.1038/s41587-024-02143-0

2. “Artificial Intelligence and Machine Learning Technology Driven Modern Drug Discovery and Development”

Sarkar, C.; Das, B.; Rawat, V.S.; Wahlang, J.B.; Nongpiur, A.; Tiewsoh, I.; Lyngdoh, N.M.; Das, D.; Bidarolli, M.; Sony, H.T.

Published online: 19 January 2023. doi.org/10.3390/ijms24032026

3. “Structure-Based Drug Discovery with Deep Learning”

R. Ozcelik, D. van Tilborg, J. Jimenez-Luna, and F. Grisoni

Published online: 13 June 2023. doi.org/10.1002/cbic.202200776

4. “Highlights on Future Treatments of IPF: Clues and Pitfalls.”

Libra, A.; Sciacca, E.; Muscato, G.; Sambataro, G.; Spicuzza, L.; Vancheri, C.

Published online: 01 August 2024. doi.org/10.3390/ijms25158392

5. “PandaOmics: An AI-Driven Platform for Therapeutic Target and Biomarker Discovery”

Petrina Kamyra, Ivan V. Ozerov, Frank W. Pun, Kyle Tretina, Tatyana Fokina, Shan Chen, Vladimir Naumov, Xi Long, Sha Lin, Mikhail Korzinkin, Daniil Polykovskiy, Alex Aliper, Feng Ren, and Alex Zhavoronkov

Published online: 26 February 2024. doi.org/10.1021/acs.jcim.3c01619

6. “Chemistry42: An AI-Driven Platform for Molecular Design and Optimization”

Yan A. Ivanenkov, Daniil Polykovskiy, Dmitry Bezrukov, Bogdan Zagribelnyy, Vladimir Aladinskiy, Petrina Kamyra, Alex Aliper, Feng Ren, and Alex Zhavoronkov

Published online: 02 February 2023. doi.org/10.1021/acs.jcim.2c01191

7. “Artificial intelligence integration in the drug lifecycle and in regulatory science: policy implications, challenges and opportunities.”

Oualikene-Gonin W, Jaulent M-C, Thierry J-P, Oliveira-Martins S, Belgodère L, Maison P, Ankri J and The Scientific Advisory Board of ANSM

Published online: 02 August 2024. doi: 10.3389/fphar.2024.1437167