



MINISTERO DELLO SVILUPPO ECONOMICO  
DIREZIONE GENERALE PER LA TUTELA DELLA PROPRIETA' INDUSTRIALE  
UFFICIO ITALIANO BREVETTI E MARCHI

# UIBM

<b>DOMANDA NUMERO</b>	<b>102014902316899</b>
<b>Data Deposito</b>	<b>16/12/2014</b>
<b>Data Pubblicazione</b>	<b>16/06/2016</b>

<b>Sezione</b>	<b>Classe</b>	<b>Sottoclasse</b>	<b>Gruppo</b>	<b>Sottogruppo</b>
G	06	F		

Titolo

APPARATO DI DIAGNOSI DIFFERENZIALE IN MEDICINA ADATTATO PER DETERMINARE LA SEQUENZA DI TEST OTTIMALE ATTA AD IDENTIFICARE UNA PATOLOGIA SECONDO CRITERI DI APPROPRIATEZZA DIAGNOSTICA

Descrizione dell'Invenzione Industriale dal titolo:

**APPARATO DI DIAGNOSI DIFFERENZIALE IN MEDICINA ADATTATO PER DETERMINARE LA SEQUENZA DI TEST OTTIMALE ATTA AD IDENTIFICARE UNA PATOLOGIA SECONDO CRITERI DI APPROPRIATEZZA DIAGNOSTICA**

Di **L.I.A. di Giuseppe Capasso**, di nazionalità Italiana, con sede in Via F.lli Cairoli 38, 09134 Cagliari ed elettivamente domiciliata presso i Mandatari Ing. Corrado Borsano (No. Iscr. Albo 446 BM), Ing. Marco Camolese (No. Iscr. Albo 882 BM), Ing. Matteo Baroni (No. Iscr. Albo 1064 BM), Ing. Giovanni Zelioli (No. Iscr. Albo 1536 B) c/o Metroconsult S.r.l., Foro Buonaparte 51, 20121 Milano.

Inventori designati:

- **CONTU Licinio**, di nazionalità italiana, Via Argonne 9, 09100 Cagliari
- **MASIA Alessandra**, di nazionalità italiana, Via Giudice Costantino 11, 09131 Cagliari
- **CHISTOLINI Pietro**, di nazionalità italiana, Via Etruria 12, 00183 Roma
- **FRUSTAGLI Gianluca**, Via Pietrabbondante 15, 00132 Roma

Depositata il \_\_\_\_\_ No. \_\_\_\_\_

**DESCRIZIONE**

Campo di applicazione dell'invenzione

La presente invenzione si riferisce al campo degli apparati di diagnosi medica, e più precisamente a \_un apparato di diagnosi differenziale adattato per determinare la sequenza di test diagnostici atta a identificare una patologia, ottimizzando gli indici di appropriatezza diagnostica.

Stato della tecnica

La diagnosi medica è il risultato di una serie di valutazioni empiriche di probabilità basate sull'esperienza professionale del medico, sulla sua preparazione, sulla sua capacità di ragionamento clinico e intuizione e sulle conoscenze che egli possiede delle persone, dei gruppi familiari e della popolazione in cui opera, relativamente alla frequenza delle diverse patologie, alla loro eventuale aggregazione familiare, alle condizioni sociali e allo stile di vita delle persone, ed ai fattori ambientali in gioco.

In base a tutto questo e alle risultanze del colloquio col malato, alla raccolta dei dati anamnestici, dei sintomi lamentati dal paziente e dei segni clinici rilevati con l'esame fisico, il medico potrà già formulare, in un piccolo numero di casi, una diagnosi certa o

altamente probabile.

Nella maggior parte dei casi, dovrà limitarsi a formulare una o più ipotesi diagnostiche che dovrà sottoporre a verifica attraverso l'esecuzione di una serie di test diagnostici, di laboratorio e strumentali. Tale processo viene definito come diagnosi differenziale.

Il punto chiave del processo è che, benché una diagnosi possa essere corretta, non è detto che sia appropriata (se per esempio per arrivare a formulare la diagnosi si sia fatto ricorso a test diagnostici ridondanti, non necessari, ingiustificatamente costosi o inutilmente rischiosi per il paziente).

I test diagnostici oggi disponibili nella pratica medica sono numerosi e in continua evoluzione e la scelta da parte del medico per uno specifico caso può essere molto complessa. Appare evidente come la prescrizione di uno o più test possa diventare particolarmente critica per i medici di medicina generale (MMG) che si trovano a operare trasversalmente in diversi campi medici e che sono spesso il primo contatto per un paziente.

I principali rischi associati alla scelta dei test sono essenzialmente due: 1) la sovrautilizzazione di test diagnostici con conseguente aumento ingiustificato dei costi per il Sistema Sanitario Nazionale (SSN) e delle ripercussioni negative sulla salute del paziente e 2) la sottoutilizzazione, dovuta a una limitata conoscenza di un'offerta troppo vasta, che può portare a una mancata diagnosi.

Questo stato di fatto ha portato in anni recenti diverse organizzazioni medico/sanitarie nazionali e internazionali a identificare dei criteri per assicurare la qualità delle prestazioni (prescrizione di test diagnostici e trattamenti farmacologici) e il contenimento dei costi, sulla base soprattutto della valutazione degli esiti (evidence-based medicine), ovvero sia di quanto una prestazione sia efficace apportando un beneficio al paziente. In questo contesto vale la pena citare lo studio svolto dalla commissione su Quality of Health Care della National Academy of Science negli Stati Uniti (in Crossing the quality chasm: a new health system for the 21st century. Washington:National Academy Press; 2001) che identifica quali obiettivi del Sistema Nazionale del XXI secolo la sicurezza, l'efficacia, la tempestività, l'efficienza, l'equità e la centralità del paziente.

È in questo contesto che matura il concetto di appropriatezza diagnostica di recente introduzione e di sempre più ampia diffusione in campo sanitario.

L'appropriatezza di un test diagnostico per una data patologia non è determinata solo dalla

sua accuratezza, esprimibile in termini di sensibilità e specificità, ma da diversi altri fattori: i costi, i tempi di attesa per l'accesso e per il risultato, la sicurezza e la tollerabilità per il paziente, come ad esempio descritto in: Ferrante di Ruffano L., Hyde C.J., McCaffery K.J. Bossuyt P.M.M. and Deeks J.J., "Assessing the value of diagnostic tests: a framework for designing and evaluating trials", British Medical Journal. 2012; 344-352.

Il concetto di appropriatezza si estende all'intero percorso diagnostico. Infatti, lo scopo del medico è quello di individuare la malattia che con maggior probabilità affligge il paziente a partire dalla raccolta dei dati anamnestici, dei sintomi e segni, e attraverso la scelta dei test necessari per la diagnosi differenziale, secondo un percorso diagnostico accurato, efficace, rapido, sicuro ed economico, in una parola appropriato.

Per questo obiettivo è evidente come le capacità del medico potrebbero essere potenziate se avesse a disposizione diverse proposte di percorsi diagnostici appropriati tra cui scegliere e potesse sapere a priori quanto la prescrizione di uno o più test affinerebbe la diagnosi differenziale in modo efficace in termini di costi/benefici per il paziente e per il SSN.

In generale i sistemi automatici e computerizzati di supporto alla diagnosi clinica si fondano su una base di conoscenza generale (dati epidemiologici, interazioni tra farmaci, linee guida...) che va integrata con i dati specifici del paziente (anagrafica, anamnesi, sintomatologia, segni). I dati vengono quindi elaborati al fine di formulare ipotesi diagnostiche attraverso l'impiego di teorie e tecnologie dell'informazione (simulazioni, algoritmi bioinformatici e procedure statistiche, sistemi di intelligenza artificiale).

Nel passato sono stati fatti diversi tentativi di sviluppare dei software, o più in generale sistemi ICT, che potessero aiutare i medici a prendere decisioni diagnostiche. Inizialmente si è trattato di sistemi esperti con database strutturati sviluppati all'interno di strutture ospedaliere e che avevano la funzione di grandi opere di consultazione. Sistemi di vero e proprio supporto alla diagnosi, in alcuni casi con un'indicazione della probabilità dell'ipotesi diagnostica, sono stati principalmente realizzati per singole specialità mediche, ambito in cui il problema teorico e tecnico relativo alla diagnosi è a un livello di complessità contenuto. Alcuni sistemi utilizzano regole per l'inferenza (del tipo if ,then) anziché metodi probabilistici o algoritmi euristici.

Più recentemente sono stati sviluppati sistemi di supporto alla diagnosi che utilizzano una base di conoscenza non strutturata e complessi algoritmi per l'inferenza, come ad esempio

descritti nell'articolo di Waghlikar, K. B.; Sundararajan, V. & Deshpande, A. W. . "Modeling Paradigms for Medical Diagnostic Decision Support: A Survey and Future Directions", *J. Medical Systems*. 2012; 36 (5): 3029-49, oppure nell'articolo di El-Kareh R., Hasan O., Schiff G.D.. Use of health information technology to reduce diagnostic errors. *BMJ Quality & Safety*. 2013; 22(Suppl 2):ii40-ii51.

Di seguito vengono brevemente descritti a titolo esemplificativo i sistemi di supporto alla diagnosi noti di maggiore interesse clinico, dedicati a un'ampia gamma di malattie.

Un primo sistema di supporto alla diagnosi, noto con l'acronimo **DXplain**, è descritto nell'articolo di Barnett G.O., Cimino J.J., Hupp J.A. Hoffer E.P., "DXPlain - An evolving diagnostic decision-support system", *JAMA*, 1987; 258(1): pag. 67-74. Si tratta di un sistema di supporto alla decisione clinica che fornisce una lista di possibili diagnosi a partire da un certo insieme di segni, sintomi dati epidemiologici e test diagnostici (dati-paziente), mediante l'impiego della logica bayesiana. Il ranking di ogni patologia è definito da quanto l'insieme dei dati-paziente si avvicina all'insieme delle manifestazioni cliniche caratteristiche di quella malattia, attraverso tecniche di "pattern-matching". Ogni associazione malattia/dato-paziente è descritta da due numeri: 1; la frequenza, organizzata in sette livelli, con cui lo specifico dato-paziente si presenta nella malattia (legato al concetto di sensibilità) e 2) il potere evocativo, espresso su otto livelli, di un dato-paziente, ovvero con quanta forza la sua presenza conferma la malattia (in analogia al valore predittivo positivo di un test). Inoltre, ogni dato-paziente ha un valore di importanza su cinque livelli indipendente dalle malattie che indica quanto sia presente in soggetti malati o quanto raramente sia rilevabile in soggetti sani. Infine, ad ogni malattia sono associati dei valori intrinseci, uno per la prevalenza e uno per l'importanza (intesa come impatto negativo che avrebbe escluderla dalla lista).

Su richiesta il sistema fornisce una spiegazione del perché ognuna delle malattie debba essere considerata, suggerisce quale ulteriore informazione clinica sarebbe necessaria per chiarire la diagnosi differenziale, indica quali manifestazioni cliniche siano atipiche per ogni malattia e fornisce dieci riferimenti bibliografici per ogni malattia. Questo sistema è probabilmente il sistema di prima generazione più evoluto e completo.

È pure noto un sistema denominato **Medical diagnosis system** (autore G. Fiore), un software diagnostico-differenziale, per piattaforma Windows su PC e dispositivi mobili. Si tratta di un sistema di supporto diagnostico di prima generazione, simile al DXplain, in cui

ad ogni sintomo è abbinato un valore empirico (da 1 a 10) sulla base della frequenza del sintomo nelle diverse patologie (se è presente in tante malattie è più importante) e importanza con cui questo sintomo si presenta nelle diverse patologie. L'utente può fissare una soglia per la percentuale minima di considerazione di un'ipotesi diagnostica e può operare delle associazioni personalizzate tra sintomi e diagnosi. Il sistema può essere utilizzato per la consultazione attraverso una ricerca avanzata per i sintomi (del tipo: comincia con, contiene, and/or), presentati in ordine alfabetico o raggruppati per organo o per tipo (esami di laboratorio, esame obiettivo). Il sistema quindi fornisce un elenco di diagnosi possibili in ordine alfabetico. L'utente può infine scegliere attivamente se avere una stima della probabilità associata alle ipotesi diagnostiche per una diagnosi differenziale.

Recentemente sono stati proposti dei sistemi che consentono un'interrogazione rapida e facile, di una base di conoscenza non strutturata costituita da una ampia gamma di informazioni cliniche, diagnostiche, scientifiche e epidemiologiche per affiancare il medico durante la sua attività lavorativa. In tale contesto si individuano principalmente due sistemi: ISABEL (Isabel Healthcare Inc.) e WATSON (IBM).

Il Sistema **Isabel**, descritto nell'articolo di Ramnarayan P. Tomlinson A, Rao A., Coren M., Winrow A., Britto J. ISABEL "A web-based differential diagnostic aid for paediatrics: results from an initial performance evaluation", *Archives of Disease in Childhood*, 2003;88:408-413, è una lista di controllo diagnostico web-based che consente l'immissione manuale dei dati-paziente (segni-sintomi, esami di laboratorio e strumentali, dati demografici) e l'integrazione con diversi formati statunitensi del Fascicolo Sanitario Elettronico (FSE). Isabel impiega un motore di ricerca basato sull'elaborazione del linguaggio naturale per identificare le corrispondenze tra i dati-paziente immessi in forma di testo, per esempio nelle note del FSE, con termini simili nel dataset diagnostico semi-codificato. Le ipotesi diagnostiche sono ordinate in base alla forza della corrispondenza con i dati immessi. Oltre alle diverse ipotesi diagnostiche propone anche possibili terapie in base alla letteratura scientifica e ai protocolli noti. Gli algoritmi che consentono l'estrazione delle ipotesi diagnostiche sono proprietari e si basano su un database di probabilità.

Il Sistema **Watson**, descritto nella domanda di brevetto WO2012/122198-A1, è un calcolatore massivamente parallelo composto di 90 server, 2880 processori (cores), 16 Terabytes di RAM in grado di acquisire ed interpretare un input vocale (la domanda), elaborare informazioni alla velocità di 500 Gigabytes al secondo, fornire una risposta in

sintesi vocale al più entro tre secondi. In anni recenti l'IBM si è orientato alle applicazioni in ambito sanitario e sta adattando il sistema Watson per realizzare un sistema di diagnosi automatica. L'impostazione rimane la stessa descritta in precedenza. Si tratta sempre di una "QA machine" (question answering machine) che, in risposta ad una specifica domanda (il quadro clinico di un generico paziente: segni, sintomi, ecc.) fornisce in uscita l'elenco delle patologie più probabili a partire da una base di conoscenze mediche testuali immesse in memoria (testi specialistici, manuali, articoli scientifici, linee guida, ecc.). Le prime applicazioni sperimentali sono nel campo della didattica medica e in quello specialistico dell'oncologia polmonare.

A fronte dei numerosi tentativi di sviluppare sistemi di supporto alla diagnosi, molti di questi sistemi si sono dimostrati validi ma nessuno è ancora entrato nell'uso corrente nella pratica medica.

Tra i motivi di criticità dei sistemi noti, come ad esempio descritti nell'articolo di Kawamoto K., Houlihan C.A., Balas E.A., Lobach D.F. *"Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success"*, BMJ. 2005; p. 330-337, e pure descritti in Wagholikar et al, su citato, si evidenziano i seguenti:

- 1) l'effettiva correttezza della diagnosi proposta, valutabile in rapporto all'esito per il paziente;
- 2) la necessità di avere un sistema davvero integrato con l'attività corrente del medico, che sia quindi disponibile nel momento e nel luogo della visita medica;
- 3) la necessità di avere un sistema che sia facilmente e rapidamente consultabile e non richieda un lungo training;
- 4) la necessità di avere un sistema che offra delle raccomandazioni anziché delle valutazioni.

Il primo punto critico dei sistemi di supporto alla diagnosi noti risiede nell'implementazione del modello di ragionamento diagnostico del medico che non può essere ridotto né a un sistema di regole empiriche che associ i dati-paziente alle malattie, come nei sistemi di prima generazione, né a un sistema automatico del tipo "a domanda risponde".

In sintesi i criteri adottati da Watson, ma anche da Isabel e dai principali dispositivi noti, si basano sul fatto di mettere a punto una sorta di raffinato motore di ricerca (del tipo di

quelli di google o yahoo) che:

1. acceda ad una mole immensa di documentazione (medica in questo caso), codificata nel linguaggio naturale (quasi sempre in inglese);
2. utilizzi degli algoritmi di analisi del linguaggio naturale e
3. valuti tutta questa documentazione a partire dalla specifica domanda posta al dispositivo, assegnando un "opportuno" punteggio (score) a ciascun documento;
4. utilizzi un motore inferenziale (per es. una rete bayesiana) per fornire in uscita l'elenco delle diagnosi più probabili.

Questo genere di architettura pone tuttavia dei problemi in termini di accuratezza. Infatti se da una parte l'uso di una base di conoscenza non strutturata di tipo testuale è di indubbia utilità perché offre una fonte di informazioni pressoché illimitata, dall'altra introduce problemi per l'estrazione di informazione utile e correttamente interpretata. I tipici problemi associati all'elaborazione del linguaggio naturale diventano infatti critici in campo medico, come ad esempio descritto nell'articolo di Nadkarni P.M., Ohno-Machado L., Chapman W.W. "*Natural language processing: an introduction*", Journal of the American Medical Informatics Association, 2011; 18: 544-551, per l'uso di un linguaggio estremamente sintetico, di abbreviazioni, di parole composte, di sinonimi. Risultano inoltre critici la segmentazione del testo in gruppi significativi e la categorizzazione di specifiche parole. Infine la precisione della risposta è legata alla quantità di testo introdotta nella specifica richiesta (*query*), basata sul linguaggio naturale.

Infatti attualmente il sistema Watson è sperimentato in un campo medico specifico e il sistema Isabel fa ricorso a una base di conoscenza mista strutturata e non strutturata. Entrambe le strategie consentono di ridurre la complessità del problema e ad ottenere diagnosi più accurate.

Sebbene i sistemi Watson e Isabel possano fornire la risposta a una singola interrogazione in pochi secondi, i tempi necessari per formulare l'ipotesi diagnostica si allungano considerevolmente perché sono richieste reiterate *query* e il processo di immissione dati in fasi successive non appare facile.

Il secondo punto critico dei sistemi attuali consiste nel fatto che essi non tengono conto dei criteri dell'appropriatezza nelle diverse fasi del processo diagnostico. Nel presente contesto il significato di appropriatezza è noto al tecnico del ramo. A nostra conoscenza gli apparati disponibili si limitano a considerare l'efficacia di un trattamento sulla base dell'evidence-

based medicine, trascurando altri aspetti fondamentali dell'appropriatezza diagnostica che include la garanzia di tempestività, contenimento dei costi, sicurezza e centralità al paziente nelle fasi che compongono il percorso diagnostico. Una fase di elevata criticità è sicuramente quella della prescrizione dei test diagnostici, la cui varietà e numerosità pone seri problemi di sovra/sotto-diagnosi, di cui si è già riferito, ovvero di inappropriatezza con evidenti ripercussioni per la salute del paziente e di costi per il SSN.

I sistemi attuali non tengono conto della necessità di un medico di ottenere un supporto per l'intero percorso diagnostico che passa attraverso la prescrizione di esami diagnostici appropriati. Il sistema Watson propone al medico quali informazioni mancanti potrebbero aumentare la confidenza della diagnosi, ma non suggerisce un percorso diagnostico.

La sola presentazione di una lista ordinata delle probabili malattie in base al quadro clinico del paziente può non essere sufficiente per il medico a formulare una diagnosi appropriata. Saranno necessari verosimilmente test diagnostici per confermare e escludere malattie. Se non si considera l'intero percorso si rischia di rimanere nell'ambito delle valutazioni soggettive ed empiriche, facendo, inoltre, ricorso sistematico a numerosi, costosi e in parte inutili test diagnostici o trascurandone altri più appropriati, ma meno noti.

Vi è quindi la necessità di trovare un sistema di supporto alla diagnosi medica che potenzi le conoscenze del medico (empowerment), integrando le sue proprie con quelle della letteratura scientifica e delle linee guida nazionali e internazionali disponibili in un database dinamico e possa compensare la tendenza alla sotto/sovra diagnosi attraverso una scelta guidata di test secondo i criteri dell'appropriatezza diagnostica e ridurre l'esitazione insita nel giudizio umano del medico.

### Sommario dell'invenzione

Pertanto scopo della presente invenzione è quello proporre un apparato di diagnosi medica adattato per determinare la sequenza ottimale di test diagnostici atta a identificare una patologia secondo il criterio dell'appropriatezza volto a superare tutti gli inconvenienti suddetti, che sia atto a formulare percorsi diagnostici, completi di test, in base a indici di appropriatezza e che sia facilmente e rapidamente consultabile nel corso di una visita medica.

Oggetto della presente invenzione è, pertanto, un apparato di diagnosi medica di ausilio al medico durante la diagnosi differenziale, che consente di effettuare una scelta guidata dei

test diagnostici, determinando percorsi diagnostici appropriati.

Nel seguito vengono definiti come **test diagnostici** tutte le procedure codificate finalizzate a formulare una diagnosi attraverso la conferma o l'esclusione di una patologia.

A titolo esemplificativo, ma non esaustivo, sono inclusi in questa definizione gli esami di laboratorio, strumentali, clinici, e i questionari (domande standard/codificate) per il medico su segni e sintomi non rilevati e/o non riferiti dal paziente.

La sequenza dei test diagnostici determinati dall'apparato oggetto dell'invenzione viene definito **percorso diagnostico**. L'apparato oggetto dell'invenzione determina un percorso diagnostico come appropriato per formulare una diagnosi, in base alla determinazione di un indice di appropriatezza globale (dell'intero percorso diagnostico) ricavato dagli indici numerici di appropriatezza relativi a ciascun test diagnostico.

Lo scopo della presente invenzione è, quindi, quello di realizzare un dispositivo di supporto alla diagnosi medica attraverso l'elaborazione di un percorso appropriato di test diagnostici.

Quindi il dispositivo oggetto della presente invenzione è in grado di "supportare" il medico suggerendo percorsi diagnostici ottimizzati in termini di appropriatezza (per esempio il percorso meno costoso e/o più rapido).

L'apparato della presente invenzione è composto da **moduli**, essenzialmente realizzati con software implementati in dispositivi di memorizzazione leggibili ed eseguibili da un computer corredato da dispositivi di input/output.

Tali moduli sono rappresentati dallo schema a blocchi di figura 1, nel quale si distinguono due moduli principali. Un modulo **MI** svolgente la funzione di "Motore Inferenziale" ed un modulo **ePDA** dedicato alla "Elaborazione del Percorso Diagnostico Appropriato".

Il modulo **ePDA** riceve, come input dal modulo **DP**, dati relativi alle informazioni del paziente e riceve dal modulo **ePDA** i dati relativi alla specifica configurazione delle variabili di sistema da utilizzare per l'elaborazione del Percorso Diagnostico Appropriato (**PDA**). A partire da questi dati, il motore inferenziale **MI** interroga una base di conoscenza (**BDC**), elabora e fornisce al modulo **ePDA** le informazioni necessarie per elaborare i percorsi diagnostici e generare come risultato il **percorso diagnostico** che potrà essere utilizzato dal medico.

Più in particolare e sempre con riferimento alla figura 1:

Nel modulo **DP** (**DATI PAZIENTE**) vengono gestiti i dati del paziente, sia in modo

interattivo che da archivio elettronico (per es. database pazienti, fascicolo sanitario elettronico, ecc.), sia localmente che in modalità remota.

Nel modulo **cPDA** (CONFIGURAZIONE **PDA**) viene memorizzata la configurazione delle variabili di sistema includendo, in queste ultime, anche la configurazione dei parametri di appropriatezza (come, ad esempio, ma non solo, il costo, il tempo, e la sicurezza e la tollerabilità del test per il paziente) per calcolare l'indice di appropriatezza globale da ottimizzare nel percorso diagnostico.

Il modulo **BDC** (BASE DI CONOSCENZA) comprende una base di dati relazionale contenente tutti i dati relativi alle patologie, segni, sintomi, farmaci, valori di riferimento delle analisi cliniche, ricavate, ad esempio, ma non esclusivamente, da banche di dati epidemiologici, linee guida e letteratura scientifica. Vi sono inoltre contenuti i valori dei parametri di appropriatezza dei test prescrivibili per completare il percorso diagnostico come ad esempio, ma non esclusivamente, sensibilità, specificità, sicurezza, costo e tempi di attesa necessari per il computo dei percorsi più appropriati. La base di dati relazionale è popolata con metodologia nota in sé.

Il modulo **MI** (MOTORE INFERENZIALE) utilizza un algoritmo per calcolare, con procedimento del tipo inferenziale, a partire dai dati paziente, i valori delle probabilità per ogni ipotesi diagnostica. Il modulo inoltre propone per ogni ipotesi uno o più test diagnostici e calcola quanto una particolare sequenza di test (percorso diagnostico) influenzi la probabilità delle diverse ipotesi e associa a ciascun percorso il relativo indice di appropriatezza.

La procedura inferenziale adottata nel modulo **MI** può basarsi preferenzialmente, ma non esclusivamente su uno degli algoritmi noti dalla letteratura scientifica (vedere per es. la rassegna di Waghlikar e al. precedentemente citata). Ad esempio per l'inferenza diagnostica possono essere utilizzati algoritmi basati su relazioni di tipo "fuzzy logic" tra sintomi e malattie, i modelli di reti neurali artificiali, le reti bayesiane che combinano il ragionamento simbolico e l'approccio statistico bayesiano e le "support vector machine", un metodo di apprendimento supervisionato per la classificazione e la regressione non-lineare. Nel seguito di descrizione di dettaglio verrà descritto un motore inferenziale basato sull'approccio statistico bayesiano, come uno dei possibili algoritmi implementabili.

Il modulo **oPDA** (Output **PDA**) visualizza una lista di ipotesi diagnostiche (ossia di patologie) considerate più probabili e per ognuna di esse elabora e mostra i possibili

percorsi diagnostici con i rispettivi indici di appropriatezza.

Aspetti caratterizzanti della presente invenzione sono i seguenti:

1. L'apparato determina una lista di ipotesi diagnostiche, a partire dai dati paziente e dalla base di conoscenza, e quantifica la probabilità ad esse associata (probabilità pre-test).
2. L'apparato determina per ogni ipotesi diagnostica percorsi diagnostici appropriati utilizzando informazioni prelevate da basi di dati consolidate e aggiornate, registri medici ed epidemiologici provenienti da istituzioni di livello internazionale, nazionale e regionale.
3. L'apparato associa a ciascun test e a ciascun percorso diagnostico un indice di appropriatezza in base a linee guida nazionali e internazionali.
4. L'apparato prevede quale impatto (probabilità post-test) avrà un test diagnostico, o più in generale, un percorso diagnostico a partire dalla probabilità pre-test precedentemente determinata, prima che il test venga eseguito.

Un importante aspetto dell'invenzione consiste nell'ottenere, proprio grazie al supporto descritto durante l'intero percorso diagnostico, diagnosi più rapide, meno costose, più certe, in sintesi più **appropriate** (efficaci, efficienti, sicure) nei casi più complessi che il medico generico o internista si trova a fronteggiare.

È particolare oggetto della presente invenzione un apparato di diagnosi differenziale adattato per applicazione medica per determinare una sequenza ottimale di test diagnostici atta a identificare una patologia adottando criteri di appropriatezza diagnostica, come meglio descritto nelle rivendicazioni, che formano parte integrante della presente descrizione.

#### Breve descrizione delle figure

Ulteriori scopi e vantaggi della presente invenzione risulteranno chiari dalla descrizione particolareggiata che segue di un esempio di realizzazione della stessa (e di sue varianti), e con riferimento ai disegni annessi dati a puro titolo esplicativo e non limitativo, in cui:

nella figura 1 è indicato un diagramma a blocchi funzionali dell'apparato di diagnosi medica oggetto della presente invenzione;

nella figura 2 è mostrato un diagramma di interrelazione funzionale tra blocchi del sistema di database dell'apparato;

nella figura 3 è mostrato uno schema di interazione tra i menù interattivi dell'apparato.

#### Descrizione di dettaglio di esempi di realizzazione

Si descrivere qui di seguito un esempio di realizzazione dell'apparato oggetto dell'invenzione, facendo riferimento al diagramma di flusso di fig. 1.

Al centro di tale diagramma vi è il blocco di elaborazione del percorso diagnostico appropriato **ePDA** che, data la configurazione del sistema **cPDA**, produce l'informazione in uscita **oPDA**. Questi blocchi realizzano un flusso verticale di informazione **PDA**, che si avvale, inoltre, della connessione con il database **DB** del sistema nel quale distinguiamo altri due blocchi; il database relativo ai dati del paziente **DP** e il database della base di conoscenza **BDC**. I dati provenienti da quest'ultima vengono immessi nel flusso verticale di informazione **PDA** attraverso il blocco del motore inferenziale **MI**. Le due sezioni (**BDC** e **DP**) che costituiscono il database **DB** possono essere allocate fisicamente anche su sistemi hardware distinti. Il database **DP** può essere costituito, ad esempio, dall'archivio pazienti privato del medico che si interfaccia con il sistema di supporto alla diagnosi.

Si è detto sopra che l'insieme dei test diagnostici che in ultima analisi il medico intende prescrivere viene definito **percorso diagnostico**. Un test o un percorso diagnostico può essere definito come appropriato per formulare una diagnosi secondo un principio generale di minimizzazione del rapporto costi/benefici, in base alla determinazione di diversi **parametri di appropriatezza**, associando ad ognuno di essi un valore numerico.

Esempi non limitativi di **parametri di appropriatezza** sono i seguenti:

Il primo è senza dubbio l'**Efficacia (E)** di un test nel confermare o escludere una specifica patologia. L'efficacia di un test informa di quanto la conoscenza per una patologia progredisce o regredisce in base al risultato valutabile in termini di variazione della probabilità di un'ipotesi diagnostica da prima (probabilità pre-test) a dopo l'esecuzione del test (probabilità post-test).

Per determinare se l'esecuzione di un test sia appropriata sicuramente è necessario valutarne anche il costo economico. Il parametro **Costo (C)** si riferisce al costo per il Sistema Sanitario Nazionale stabilito dal Nomenclatore Tariffario Nazionale.

Anche il **tempo (Te)** necessario per accedere al test è un parametro per definire l'appropriatezza di un test, in quanto è chiaramente necessario eseguire il test tanto più rapidamente quanto più grave è la malattia sospettata. I dati sui tempi di attesa sono quelli

riportati dal Piano Nazionale di Governo delle Liste di Attesa .

Altri aspetti che caratterizzano l'appropriatezza dei test sono la sicurezza e la tranquillità del paziente.

Il parametro **Rischio (R)** tiene conto del *Rischio intrinseco*, relativo alla procedura in sé, e il *Rischio relativo*, dipendente dalla condizione patologica del paziente.

La tollerabilità da parte del paziente è importante per garantire che la prescrizione verrà seguita e il percorso effettivamente completato. Il parametro **Tollerabilità (To)** è definito in modo soggettivo attraverso, ad esempio, un questionario da sottoporre al paziente.

Tutti i parametri di appropriatezza possono essere opportunamente ricondotti a valori numerici.

## **IL DATABASE**

Le informazioni contenute nel database **DB**, sotto forma di tabelle di record collegate da una rete di relazioni, sono essenzialmente di due tipi, come indicato genericamente nella figura 1, e nel seguito descritti in dettaglio con riferimento alla figura 2.

1. **BDC**; un articolato insieme di strutture dati precaricate che costituiscono la “**base di conoscenza**”, sulla quale l'applicazione basa tutte le scelte logiche e le elaborazioni;
2. **DP**; i dati specifici gestiti dagli utenti medici e relativi a segni, sintomi e anamnesi, inseriti nel corso delle visite mediche (**Dati Pazienti**). In **DP** ci sono, inoltre, i dati relativi ai risultati dell'elaborazione, in termini di ipotesi diagnostiche ed esami da effettuare (cioè i dati prodotti dal flusso verticale dell'elaborazione del Percorso Diagnostico Appropriato **PDA**).

### **Base di conoscenza BDC dell'applicazione.**

La base di conoscenza è strutturata come insieme di tabelle variamente interconnesse, con contenuti precaricati e aggiornabili, come da fig. 2.

La tabella **Malattia** ha un ruolo centrale nella “base di conoscenza” dell'applicazione: essa individua le malattie conosciute dall'applicazione come combinazioni univoche di categorie e sub-categorie nella classificazione **ICD10**. Infatti le tabelle collegate ad essa **CtgMltTmt\_ICD10** (Categorie Malattie e Traumatismi ICD10) e **SubctgMltTmt\_ICD10** (Subcategorie Malattie e Traumatismi ICD10) contengono appunto i corrispondenti elenchi della decima revisione della classificazione internazionale di malattie e problemi correlati

proposta dall'OMS (ICD10).

Alla tabella Malattia è collegata la tabella **Prevalenza** che riporta il valore di probabilità per ognuna delle malattie, normalizzata per sesso, etnia e range di età del paziente.

Le tabelle **Sintomo**, **Segno** e **Test** contengono, rispettivamente, l'elenco di tutti i sintomi e segni riscontrabili e l'elenco di tutti i test medici (test strumentali) effettuabili. Queste tabelle sono ciò che permette all'applicazione di interpretare le informazioni immesse dal medico durante le visite, permettendo di scegliere fra le varie possibilità previste.

La tabella **Test**, oltre al nome e alla descrizione estesa del test, riporta anche se si tratta di un test dicotomico (positivo/negativo – campo: Dicotomico) o a valori continui (e, nel caso sia del tipo a valori continui, anche l'unità di misura utilizzata – campo: Unita).

Ciò è di fondamentale importanza per l'algoritmo di calcolo in quanto si è scelto come approccio quello di trattare logicamente e ridurre ogni test a valori continui a una serie di test dicotomici “virtuali” suddividendo l'intervallo di variabilità dei possibili risultati in un adeguato numero di sottointervalli (secondo questo approccio ogni valore risultante dall'esecuzione del test cadrà all'interno di uno dei possibili sottointervalli rendendo positivo il sotto-test dicotomico corrispondente).

Questa tabella comprende anche i valori numerici di alcuni parametri di appropriatezza associati al test: costo, tempo di attesa e rischio intrinseco.

La tabella **Sintomo\_Malattia** esplicita la relazione multi-a-molti esistente tra la tabella Sintomo (ogni malattia può avere più sintomi -o nessuno- e ogni sintomo può essere comune a più malattie) e la tabella Malattia e include come attributi la probabilità di presenza del sintomo nei soggetti affetti dalla corrispondente malattia e la probabilità di assenza del sintomo nei soggetti non affetti dalla corrispondente malattia.

Analogamente alla precedente, la tabella **Segno\_Malattia** esplicita la relazione multi-a-molti tra la tabella Segno e la tabella Malattia.

Anche la tabella **Test\_Malattia** esplicita la relazione multi-a-molti esistente tra la tabella Test (i test strumentali) e la tabella Malattia (ogni malattia potrebbe essere individuata da uno o più test combinati o in alternativa -o nessuno- e ogni test potrebbe essere utile all'individuazione di più malattie). In questo caso gli attributi sono la sensibilità e la specificità del test rispetto alla specifica malattia.

La tabella **FattoreRischio** contiene una lista di fattori di rischio di vario tipo che possono predisporre o causare malattie.

Una relazione multi-a-molti lega i fattori di rischio alla tabella Malattia, tale relazione è implementata dalla tabella **FattoreRischio\_Malattia**.

Le tabelle **Professione** ed **Etnia** contengono una lista rispettivamente di professioni ed etnie riconosciute e rilevanti dal punto di vista medico. Sono entrambe utilizzate (attraverso relazioni 1-a-molti) dalla tabella Paziente. La tabella Etnia è relazionata anche alla tabella Prevalenza.

### **Table e strutture per i dati specifici del paziente DP.**

I dati specifici del paziente prevedono un insieme di tabelle variamente interconnesse, con contenuti precaricati e aggiornabili, come da fig. 3.

L'applicazione è ovviamente costruita per essere utilizzata da medici diversi, quindi la tabella **Medico** contiene alcune informazioni anagrafiche e tutti i dati specifici relativi a ciascuno dei pazienti gestiti ed inseriti nel corso del tempo verranno memorizzati nel database e relazionati all'identificativo del medico.

La tabella **Paziente** contiene le informazioni di base per ciascuno dei pazienti.

La tabella Paziente è collegata, da una relazione 1-a-molti, alla tabella **Episodio**. Ogni record di tale tabella identifica un *episodio* nella storia clinica del paziente che inizia dalla prima visita durante la quale il medico raccoglie segni e sintomi relativi a un certo disturbo e si conclude con la diagnosi, dopo eventuali esami strumentali. Nel caso che il paziente si ripresenti, successivamente, al medico, per un altro disturbo (diverso o analogo al precedente), il medico, utilizzando l'applicazione, determinerà l'apertura di un altro episodio clinico che comporterà la creazione di un nuovo record nella tabella Episodio, collegato al medesimo paziente.

La tabella Episodio è relazionata alle tabelle **Episodio\_Segno** ed **Episodio\_Sintomo**: durante la visita il medico inserisce i sintomi riportati dal paziente. Le due tabelle Episodio\_Segno ed Episodio\_Sintomo esplicitano, popolandole, due relazioni multi-a-molti fra Episodio e Segno ed Episodio e Sintomo.

Con la tabella **Anamnesi** avviene qualcosa di analogo. La tabella Anamnesi è collegata alla tabella Paziente. I dati anamnestici del paziente vengono rilevati una sola volta ed eventualmente aggiornati nel tempo. Anche la tabella Anamnesi esplicita una relazione multi-a-molti fra la tabella Paziente e la tabella FattoreRischio. I dati anamnestici vengono raccolti elencando tutti i fattori di rischio a cui è esposto il paziente.

Dopo l'inserimento di tutte le predette informazioni l'applicazione è pronta alla prima

elaborazione per il calcolo delle prime ipotesi diagnostiche su base probabilistica, (probabilità pre-test) mediante l'uso dell'algoritmo implementato nel motore inferenziale MI.

La lista di ipotesi diagnostiche pre-test risultante viene quindi memorizzata nella tabella **Diagnosi**.

Successivamente si passa al calcolo delle probabilità post-test mediante il motore inferenziale **MI** che fa ricorso ad una serie di query che incrociano la lista delle ipotesi diagnostiche pre-test in **Diagnosi** con le tabelle **Test\_Malattia** e **Test** della base di conoscenza, ricavando per ognuna delle malattie ipotizzate tutti i test clinici strumentali associati e i rispettivi valori di sensibilità, specificità e i parametri per il calcolo del livello di appropriatezza di ogni test.

Si determina così quali fra i test individuati sono più utili a guadagnare informazione per la diagnosi finale.

Questi risultati vengono memorizzati nella tabella **Test\_Diagnosi**, dove a ogni ipotesi diagnostica sono associati una serie di test con i relativi valori dei parametri intrinseci di appropriatezza (costo, tempo di attesa, rischio intrinseco) e le probabilità post-test.

Nella stessa tabella possono essere successivamente inseriti i valori di ulteriori parametri di appropriatezza come il rischio relativo del test e tollerabilità utilizzati nella fase di ottimizzazione che prevede il calcolo di un indice di appropriatezza del test per fornire al medico una informazione in più per la scelta finale dei test da prescrivere.

In questi stessi record possono anche trovare successivamente posto i valori dei risultati reali dei test diagnostici effettivamente eseguiti a seguito della prescrizione per un loro futuro utilizzo.

### **MOTORE INFERENZIALE MI.**

È il modulo nel quale viene implementato l'algoritmo per calcolare, a partire dai dati paziente, i valori delle probabilità per ogni ipotesi diagnostica. Il modulo inoltre propone per ogni ipotesi uno o più test diagnostici e calcola quanto un test ed una particolare sequenza di test influenzi la probabilità delle diverse ipotesi e associa a ciascun percorso il relativo indice di appropriatezza.

La procedura inferenziale può basarsi, ad esempio, su uno dei qualsiasi algoritmi citati nella letteratura scientifica (tra i più comuni si possono citare quelli basati su relazioni di *fuzzy logic* tra sintomi e malattie, le reti neurali artificiali, le reti bayesiane e le *support*

*vector machine*). Nella descrizione dettagliata verrà presentato e descritto un motore inferenziale basato su un approccio bayesiano noto come NAIVE BAYES come uno dei possibili algoritmi implementabili. Più in generale la scelta dell'algoritmo non porta a diagnosi finali differenti ma può influire sul percorso diagnostico in termini di maggiore o minore velocità nel pervenire alla diagnosi finale. In altri termini ciascun algoritmo, anche per effetto dell'interattività della procedura, porterà comunque alla stessa diagnosi anche se, eventualmente, con tempi diversi.

La scelta dell'algoritmo determinerà, caso per caso, la necessità di modificare il **DB** per includervi quei dati necessari per l'esecuzione dell'algoritmo stesso.

In ogni caso il **DB** come sopra descritto, possiede già le informazioni per implementare un algoritmo di tipo NAIVE BAYES.

Consideriamo ora, a titolo esemplificativo, il caso più generale di una rete bayesiana. Matematicamente, una rete bayesiana è un grafo aciclico orientato in cui i nodi rappresentano le variabili, gli archi rappresentano le relazioni di dipendenza statistica tra le variabili e le distribuzioni locali di probabilità dei nodi foglia rispetto ai valori dei nodi padre.

In generale ogni patologia, segno, sintomo, dato anamnestico, test è, indistintamente, una variabile statistica e quindi un nodo della nostra rete bayesiana. Ad ogni arco orientato che lega un generico nodo A con un nodo B deve essere associata una probabilità condizionata  $P(B|A)$ ; la conoscenza dei valori di probabilità di tutti gli archi della rete consente di poter calcolare formalmente la probabilità per un nodo qualsiasi condizionata alla probabilità associata ad un qualsiasi altro insieme di altri nodi dati. Saremo così formalmente in grado di calcolare le probabilità  $Pre(M | \{DP\})$  e  $Post(M | \{DP\} \cup \{T\})$  cioè le probabilità pre e post-test di riscontrare la patologia M condizionata rispettivamente all'insieme dei dati paziente  $\{DP\}$  e all'unione di  $\{DP\}$  con l'insieme dei test  $\{T\}$ .

Per far ciò dovremo aggiungere al **DB** precedentemente descritto le informazioni relative alla struttura della rete (una tabella multi-a-molti) e il valore delle funzioni di probabilità associate ad ogni arco.

Quindi, come accennato, in precedenza, la scelta dell'algoritmo comporta delle aggiunte alla struttura del database. A questo proposito accenniamo ad un altro tipo di implementazione del motore inferenziale MI.

In precedenza abbiamo citato, tra gli altri, gli algoritmi denominati *support vector*

*machine*; questi ultimi sono un caso particolare delle cosiddette tecniche di *machine learning*. A titolo esemplificativo tali tecniche possono essere implementate nel seguente modo: la probabilità di una patologia viene rappresentata come una funzione, non lineare, dei dati paziente (segni, sintomi, test, ecc.). Noti questi ultimi la suddetta probabilità può essere comodamente calcolata. La difficoltà consiste nel conoscere i valori dei parametri e dei coefficienti di tale funzione che non sono noti a priori. Essi possono essere determinati ricorrendo a procedure di apprendimento tipiche delle tecniche di *machine learning*. I valori dei parametri e dei coefficienti una volta determinati, vanno inseriti in una tabella supplementare multi-a-molti del database.

Per quanto riguarda il flusso informativo verticale **PDA** di fig.1, esso comprende tre fasi funzionali successive in ordine di tempo (l'utente potrà comunque sempre tornare ad una fase precedente per rivedere e modificare le impostazioni anteriori):

1. **Configurazione Percorso Diagnostico appropriato** (blocco **cPDA** di fig. 1);
2. **Selezione delle patologie per la diagnosi differenziale** (blocco **ePDA** di fig. 1);
3. **Selezione dei Test diagnostici** (blocco **oPDA** di fig. 1).

**Configurazione Percorso Diagnostico appropriato**(blocco **cPDA** di fig. 1).

Nella fase di Configurazione Percorso Diagnostico Appropriato è consentito visualizzare, selezionare ed eventualmente modificare le variabili di sistema, ovvero i parametri, che verranno utilizzati nelle fasi successive per elaborare e generare in output i percorsi diagnostici. Tra questi parametri ci sono anche quelli necessari per calcolare gli indici di appropriatezza. In ognuna delle fasi successive alla presente si potrà sempre ritornare in questa fase per riconfigurare parzialmente o completamente l'intero apparato.

**Selezione delle patologie per la diagnosi differenziale.** (blocco **ePDA** di fig. 1).

In questa fase l'apparato si collega al database dei pazienti DP consentendo di acquisire, oltre le generalità del paziente stesso, anche i dati anamnestici necessari per le elaborazioni successive.

Al paziente così selezionato viene associata una lista di segni e sintomi riscontrati e/o riferiti al momento della visita medica.

Queste informazioni consentono di generare una Lista di Patologie più Probabili (LPP) elaborata dal motore inferenziale MI. Ad essa si unisce una Lista delle Malattie Rare (LMR), elaborata inserendo anche patologie non presenti in LPP, ma presenti nel database **DB**, per le quali, tuttavia, uno dei segni/sintomi rilevati sia altamente specifico (oppure

abbia una alta sensibilità).

La lista totale potrà sempre essere rivista ed aggiornata interattivamente dall'utente medico. In particolare, per verificare di aver considerato tutti i sintomi rilevanti per le patologie di interesse, si potrà selezionare una o più delle patologie fra quelle in lista per approfondire i sintomi correlati e verificare la presenza di ulteriori sintomi tipici ma non ancora rilevati/riferiti.

Successivamente viene elaborata interattivamente una lista delle patologie delle quali si intendono stimare i percorsi diagnostici; nella lista delle patologie probabili LPP vengono selezionate le patologie che presentano una probabilità post segni-sintomi superiori a una soglia prefissata e modificabile (in **cPDA**). Ad LPP e alla lista delle malattie rare, si può aggiungere la lista delle patologie aggiunte manualmente LPA; quindi la lista totale LPS è data da  $LPS = LPP + LMR + LPA$ .

Pertanto, in questa fase, è possibile modificare ed estendere LPP:

- Manualmente: in ogni momento è possibile aggiungere una qualsiasi patologia.
- Modificando la soglia della selezione delle patologie più probabili.
- Modificando direttamente la lista: qualsiasi patologia può essere selezionata/deselezionata dalla lista finale

Una volta consolidata LPS si può passare alla fase successiva.

**Selezione dei Test diagnostici** (blocco **oPDA** di fig. 1).

Ad ogni patologia della lista selezionata LPS corrisponde (ottenuto interrogando il DATABASE) un insieme di test. L'unione di tutti questi insiemi per tutte le patologie costituisce un insieme TAMS (test associati alle malattie selezionate).

In generale, per la selezione dei test diagnostici da eseguire, possono essere utilizzate diverse modalità o combinazioni di modalità:

1. Selezione interattiva semiautomatica dei test ad alto effetto diagnostico. Dall'insieme TAMS di tutti i test rilevanti vengono ulteriormente selezionati solamente quelli che hanno una sensibilità e/o una specificità al di sopra di una soglia di *efficacia* il cui valore può essere modificato. Successivamente viene calcolata la probabilità post-test per ogni patologia nell'ipotesi di effettuare ogni singolo test diagnostico. In questo modo vengono identificati i test che, se eseguiti, fanno progredire o regredire la conoscenza. I test ad alto effetto diagnostico saranno quelli che superano una Soglia di Probabilità Post-Test anche questa

modificabile dal medico.

2. Scelta manuale del test. Per ogni patologia il medico può visualizzare tutti i test ad essa associati e selezionare a suo piacimento uno o più test diagnostici. Può quindi accadere che anche un test che non superi la soglia di efficacia possa essere scelto. Per ogni test selezionato il sistema calcola la probabilità post-test e l'appropriatezza.
3. Se necessario il medico può anche selezionare dal **DB** un qualsiasi test indipendentemente dalla sua rilevanza. I test selezionati con questa opzione verranno inseriti nell'insieme dei TAMS.

Se viene scelto più di un test diagnostico l'apparato calcolerà anche la probabilità post-test di ogni ipotesi diagnostica, nell'ipotesi di eseguire tutti i test, e l'appropriatezza dell'intero percorso diagnostico (cioè l'appropriatezza globale di tutti i test selezionati) oltre a visualizzare la probabilità post-test e l'appropriatezza di ogni singolo test.

La sopra citata procedura associa ad ogni test il rispettivo indice di appropriatezza **IA** che, a livello esemplificativo, può essere definito come:

$$IA = T_o / (C * T_e * R)$$

dove  $T_e$  è l'indice relativo al tempo di attesa per accedere al test ;  $C$  è l'indice relativo al costo del test;  $R$  è l'indice relativo al massimo tra il rischio intrinseco e il rischio relativo del test;  $T_o$  è l'indice relativo alla tollerabilità al test.

Viene anche determinato un indice di appropriatezza globale ( $IA_G$ ) di tutti i test diagnostici selezionati, calcolata con la stessa formula precedente, vale a dire:

$$IA_G = T_{oG} / (C_G * T_{eG} * R_G)$$

dove:  $T_{oG}$  è pari al minimo tra tutti gli indici di tollerabilità tra tutti i test della sequenza,  $C_G$  è pari alla somma dei singoli costi di ciascun test della sequenza,  $T_{eG}$  è pari al massimo fra tutti i tempi di attesa,  $R_G$  è pari al massimo fra tutti gli indici di rischio della sequenza.

Ricapitolando, in uscita vengono forniti i seguenti dati, preferibilmente visualizzati su display (nel blocco output **oPDA** di fig.1) per ogni ipotesi diagnostica:

- la probabilità pre-test,
- l'elenco, in ordine di appropriatezza, di tutti i test con la rispettiva probabilità post-test ed indice di appropriatezza,
- la probabilità post-test e l'indice di appropriatezza globale dell'intero percorso diagnostico.

Si possono costruire diversi percorsi diagnostici, anche aggiungendo e togliendo singoli test dalla selezione e ricalcolare le probabilità post-test delle patologie e l'indice di appropriatezza dei percorsi.

È possibile pertanto definire un percorso alternativo (selezionando un insieme diverso di test): l'apparato mantiene in memoria il percorso precedente con i relativi valori di appropriatezza in modo tale che possano essere confrontati diversi percorsi.

#### **Descrizione di dettaglio di un esempio realizzativo.**

La presente invenzione è preferibilmente realizzata tramite moduli software memorizzati su dispositivi di memorizzazione leggibili ed eseguibili da un computer (ad esempio dei tipi server, desktop, workstation, portatile, tablet, smartphone, ecc.) corredato da dispositivi di memorizzazione di massa e di input/output (vedi schema Fig.1).

Ad esempio i moduli software sono implementati su cluster di calcolo (fisico o virtuale) ed eseguibili da remoto ed implementati come App per dispositivi mobili (Tablet, smartphone, PC portatili).

La realizzazione della parte hardware è a portata del tecnico del ramo, e non è necessario inserire informazioni di dettaglio ulteriori.

Tali moduli software utilizzeranno le informazioni memorizzate nella base di conoscenza BDC e nell'archivio dei dati dei pazienti DP ed entrambe queste basi di dati sono illustrate sopra nella sezione denominata "IL DATABASE".

Tutti i moduli software e l'intero database sono sviluppati facendo ricorso a un framework Ruby on Rails, utilizzato come ambiente RAD, corredato da una serie di moduli aggiuntivi opensource (gems) messi a disposizione dalla comunità di sviluppatori che sostiene questo ambiente. In questo genere di soluzione implementativa si prevede (nella quasi totalità dei casi) che tutti i dati e le informazioni non volatili necessarie al funzionamento dell'applicazione siano gestite mediante l'uso di un database relazionale. In questo caso si utilizza un server di database MySQL, connesso all'applicazione mediante socket locale. Il tecnico del ramo è in grado di realizzare la parte software utilizzando le sue conoscenze di base e quanto descritto nella presente descrizione.

Con riferimento alla figura 3, sul dispositivo di output verranno visualizzati una sequenza di menù interattivi di cui i principali sono tre:

- *Menù Home*
- *Menù Selezione delle patologie per la diagnosi differenziale*

- *Menù Selezione dei Test diagnostici*

Tali menù richiameranno i seguenti altri menù e le due procedure principali:

- *CPDA (Configurazione Percorso Diagnostico Appropriato)*
- *Dati Paziente*
- *Selezione segni e sintomi*
- *Lista Sintomi*
- *Selezione Sintomi Rilevanti*
- *Aggiungi Malattia Rara*
- *SIT (selezione interattiva test)*
- *SMT (selezione manuale del test)*

Procedure:

- **LP** – calcola la lista ordinata delle patologie più probabili
- **OPDA** – elabora l'Output del Percorso Diagnostico Appropriato

#### **Descrizione operativa di ogni singolo menù e delle procedure.**

(associata ad ogni opzione del menù viene riportata la descrizione dell'opzione e la lista delle attività, in pseudo codice, che vengono eseguite all'attivazione di ciascuna opzione).

#### **MENÙ HOME**

Nel menù è consentito chiamare i sottomenù nei quali configurare il sistema (menù **CPDA**), identificare il paziente e i suoi dati clinici, selezionare segni, sintomi e dati anamnestici (menù dati paziente). Queste informazioni consentono di elaborare una Lista di Patologie più Probabili (LPP) e di visualizzarla in questo menù. Ad essa si unisce la Lista delle Malattie Rare (LMR). La lista totale potrà sempre essere rivista ed aggiornata usando le opzioni del menù che vengono qui di seguito elencate.

**Opzione “sintomi ignorati per patologia”**: per verificare di aver considerato tutti i sintomi rilevanti per le patologie di interesse.

Si seleziona una delle patologie fra quelle in lista per approfondire i sintomi correlati e verificare la presenza di ulteriori sintomi tipici ma non ancora rilevati/riferiti.

Se effettivamente presenti, si possono immettere nuovi sintomi non inseriti nella prima fase di indagine. La lista delle patologie più probabili verrà di conseguenza aggiornata.

**Opzione “sintomi ignorati per importanza”**: per affinare la diagnosi differenziale fra le ipotesi più probabili.

Per le prime ipotesi diagnostiche in lista, pari alla variabile NuSePa (ad es. 5 per default),

vengono estratti tutti i sintomi non segnalati nella prima fase di indagine. Vengono visualizzati solo i primi 10 (valore di default della variabile SiPr) sintomi con più alto LR (likelihood ratio) assoluto (massimo LR di un sintomo fra le patologie più probabili).

Se effettivamente presenti, i nuovi sintomi non inseriti nella prima fase di indagine possono essere inseriti. La lista delle patologie più probabili verrà di conseguenza aggiornata.

**Opzione “sintomi già inclusi significativi”:** per verificare di non avere trascurato una patologia poco probabile (malattia rara) sebbene un sintomo altamente rilevante fosse stato rilevato.

Tra tutti i sintomi segnalati nella prima fase di indagine vengono visualizzati i primi 10 (valore di default della variabile SiPiuP) sintomi con più alto LR e la patologia per la quale tale valore è massimo.

La patologia di bassa probabilità può essere inserita nella lista delle malattie da considerare nei menù successivi.

**Menù Home (procedura).**

opzioni possibili:

**Configurazione Percorso Diagnostico Appropriato**

entra nel menù CPDA

**Dati Paziente**

entra nel menù Dati Paziente

nell'output viene visualizzato quanto segue:

IF ( “non sono ancora stati inseriti i dati del paziente e la sintomatologia”)

THEN PRINT (“inserisci i Dati Paziente, i segni e i sintomi”)

ELSE chiama la procedura **LP**

visualizza la lista LPP + LMR

attiva opzioni supplementari nel *Menù Home*

opzioni supplementari:

**Sintomi ignorati per patologia**

seleziona una patologia della lista (cliccandoci sopra)

entra nel menù Lista Sintomi

**Sintomi ignorati per importanza**

entra nel menù Selezione Sintomi Rilevanti

### **Sintomi già inclusi significativi**

entra nel menù Aggiungi Malattia Rara

### **Avanti**

entra nel menù

Selezione delle patologie per la diagnosi differenziale

## **MENÙ SELEZIONE DELLE PATOLOGIE PER LA DIAGNOSI DIFFERENZIALE**

In questo menù si elabora la lista delle patologie delle quali si intendono stimare i percorsi diagnostici.

Le tre principali opzioni sono:

**Selezione automatica:** vengono selezionate le patologie che presentano una probabilità post segni-sintomi superiori a una soglia del 40% (valore default della variabile SoSePa\_inf) ma vengono mantenute in lista le “malattie rare” eventualmente selezionate nel menù precedente.

**Selezione manuale:** il medico è libero in ogni momento di aggiungere una qualsiasi patologia

**Modifica la lista:** il medico può selezionare/deselezionare qualsiasi patologia dalla lista finale

### **Menù *Selezione delle patologie per la diagnosi differenziale* (procedura).**

opzioni possibili:

#### **Selezione automatica**

aggiorna la LPP considerando solo le patologie con un valore di probabilità maggiore o uguale a quello della variabile SoSePa\_inf

#### **Selezione manuale**

il medico inserisce una patologia che ritiene rilevante, quest'ultima viene immessa nella lista LPA (patologie aggiunte manualmente dal medico) insieme con il rispettivo valore di probabilità pre-test.

#### **Modifica la lista**

viene visualizzata l'intera Lista delle Patologie sinora Selezionate LPS;

le patologie possono essere deselezionate/selezionate

cliccandoci sopra; quelle selezionate sono visualizzate in grassetto.

### **Configurazione**

entra nel menù **CPDA**

**indietro**

torna al **Menù Home**

**avanti**

entra nel menù **Selezione dei Test diagnostici** nell'output viene visualizzata la Lista delle Patologie Selezionate (LPS) costituita dalla somma della lista LPP (aggiornata con il valore della variabile SoSePa\_inf) + LMR + la lista delle patologie aggiunte manualmente dal medico (LPA), quindi  $LPS = LPP + LMR + LPA$

**MENÙ SELEZIONE DEI TEST DIAGNOSTICI**

Ad ogni patologia della lista selezionata LPS corrisponde (ottenuto interrogando il DATABASE) un insieme di test. L'unione di tutti questi insiemi costituisce l'insieme TAMS (test associati alle malattie selezionate). Per la costruzione preliminare dell'insieme TAMS non vengono prese in considerazione tutte le patologie di LPS ma da quest'ultima lista vengono esclusi i test associati alle patologie la cui probabilità pre test è già superiore al 70% (valore di default della variabile SoSePa\_sup). Il senso di tale scelta deriva dal fatto che patologie con una alta probabilità pre-test forniscono già una forte indicazione diagnostica senza necessità di ricorrere ad altri test che aggiungerebbero una minore informazione rispetto a quella ottenibile facendo test più specifici per le patologie presenti all'interno dell'intervallo di incertezza diagnostico (che è individuato dalle variabili SoSePa\_sup e SoSePa\_inf che per default è pari al 70% e al 40%).

In generale, per la selezione dei test diagnostici da eseguire, il medico può seguire diverse modalità o combinazioni di modalità:

**Selezione interattiva semiautomatica dei test ad alto effetto diagnostico:**

Dall'insieme TAMS di tutti i test rilevanti vengono ulteriormente selezionati solamente quelli che hanno un LR sopra una soglia di *efficacia* (SoEf) il cui valore può essere fissato dal medico nel menù **Configurazione PDA** (per default SoEf = 2).

Viene calcolata la probabilità post-test per ogni patologia nell'ipotesi di effettuare ogni singolo test diagnostico. In questo modo vengono identificati i test che se eseguiti fanno progredire o regredire la conoscenza. I test ad alto effetto diagnostico saranno quelli che superano una Soglia di Probabilità Post-Test SoPPT (per default del 70% e comunque minore od uguale a SoSePa\_sup).

**Scelta manuale del test**

Per ogni patologia il medico può visualizzare tutti i test ad essa associati e ordinati secondo il rapporto di probabilità e selezionare a suo piacimento uno o più test diagnostici. Può quindi accadere che anche un test che non superi la soglia di efficacia possa essere scelto. Per ogni test selezionato il sistema calcola la probabilità post-test e l'appropriatezza.

Se necessario il medico può anche selezionare dall'elenco *Test diagnostici* un qualsiasi test indipendentemente dalla rilevanza per le ipotesi diagnostiche e dalla loro efficacia per l'analisi differenziale. I test selezionati con questa opzione verranno aggiunti nell'insieme dei TAMS.

Se viene scelto più di un test diagnostico il sistema calcolerà anche la probabilità post test nell'ipotesi di eseguire tutti i test e l'appropriatezza dell'intero percorso diagnostico oltre a visualizzare la probabilità post-test e l'appropriatezza di ogni singolo test.

Il medico può pertanto costruire a suo piacimento diversi percorsi diagnostici aggiungendo e togliendo singoli test dalla selezione e ricalcolare le probabilità post-test delle patologie e l'appropriatezza dei percorsi.

**Menù Selezione dei Test diagnostici (procedura).**

opzioni possibili:

**Selezione interattiva semiautomatica dei test ad alto effetto diagnostico**

entra nel menù SIT

**Scelta manuale del test**

viene selezionata una patologia cliccandoci sopra e si

entra nel menù SMT

**Configurazione**

entra nel menù CPDA

**Indietro**

entra nel menù Selezione delle patologie per la diagnosi differenziale

**Salva**

Memorizza la tabella del percorso diagnostico (con un nome)

**Richiama**

Carica una tabella di percorso diagnostico precedentemente salvata

**Stampa**

Stampa la tabella del percorso diagnostico in uscita viene visualizzato l'output della procedura **OPDA** (Output Percorso Diagnostico Appropriato). In dettaglio, per ogni patologia viene visualizzata la probabilità pre-test e, successivamente, vengono elencati, in ordine di appropriatezza, tutti i test con relativa probabilità post-test ed appropriatezza ed, infine, la probabilità e l'appropriatezza dell'intero percorso diagnostico.

### **MENÙ DATI PAZIENTE**

Da questo menù ci si collega al database dei pazienti consentendo, inoltre, di acquisire, oltre le generalità del paziente stesso, anche i dati anamnestici necessari per le elaborazioni successive.

Al paziente così selezionato viene associata, in questo menù, una lista di segni e sintomi riscontrati e/o riferiti al momento della visita medica.

La selezione dei segni e sintomi può avvenire in molti modi. Il modo più semplice è quello utilizzare un motore di ricerca all'interno del **DB** dei sintomi, si scrive il sintomo o il segno riscontrato in una finestra di dialogo e, a partire da questo, il motore di ricerca fornisce una lista di segni e sintomi, simili a quello immesso, dai quali selezionare quello che si considera più appropriato a descrivere la situazione del paziente.

Al termine della selezione si è in possesso di una lista di segni e sintomi che insieme alle informazioni anamnestiche del paziente (già presenti nel database) consente di tornare al menù home ed elaborare la lista delle patologie più probabili.

#### **Menù Dati Paziente (procedura).**

opzioni possibili:

#### **Selezione Paziente**

vengono fornite le generalità del paziente

#### **Selezione segni e sintomi**

si immettono i segni e i sintomi selezionandoli nel **DB**

#### **Indietro**

ritorna al Menù Home

nell'output viene visualizzato:

le generalità del paziente (una volta che sia stato selezionato)

la lista dei segni e sintomi riscontrati e selezionati nel paziente

## **MENÙ LISTA SINTOMI**

Nel menù vengono visualizzati tutti i segni e sintomi associati ad una predefinita patologia. Ne vengono visualizzati al massimo un numero pari alla variabile NuMaxSint (per default 20) e tra questi, evidenziati in grassetto (ovvero già selezionati), quelli già riscontrati dal medico. In questo modo è possibile

verificare di aver considerato tutti i sintomi rilevanti per la patologia di interesse e verificare la presenza di ulteriori sintomi tipici non ancora rilevati/riferiti.

Se effettivamente presenti, il medico può immettere i nuovi sintomi non inseriti nella prima fase di indagine. La lista delle patologie più probabili verrà di conseguenza aggiornata una volta che si ritorni nel menù Home.

### **Menù Lista Sintomi (procedura).**

opzioni possibili:

#### **Configurazione**

*entra nel menù CPDA*

#### **Ritorna**

*torna nel Menù Home*

l'output visualizza la lista di 20 (valore di default della variabile NuMaxSint) segni e sintomi per quella patologia

I segni e sintomi selezionati sono visualizzati in grassetto.

Se ne possono selezionare (deselezionare) altri cliccandoci sopra

## **MENÙ SELEZIONE SINTOMI RILEVANTI**

Con questa opzione è consentito rilevare segni e sintomi importanti ma precedentemente ignorati allo scopo di affinare la diagnosi differenziale fra le ipotesi più probabili.

Dalle prime 5 (valore di default della variabile NuSePa) ipotesi diagnostiche in lista vengono estratti tutti i sintomi non segnalati nella prima fase di indagine. Vengono visualizzati solo i primi 5 (valore di default di SiPr) sintomi con più alto LR (Likelihood Ratio) assoluto (massimo LR di un sintomo fra le patologie più probabili).

Se effettivamente presenti, il medico può immettere i nuovi sintomi non inseriti nella prima fase di indagine. La lista delle patologie più probabili verrà di conseguenza aggiornata una volta che si ritorni nel menù Home.

### **Menù Selezione Sintomi Rilevanti (procedura).**

opzioni possibili:

**Configurazione**

*entra nel menù CPDA*

**Ritorna**

*torna nel Menù Home*

l'output visualizza una lista al più di NuSePa \* SiPr segni e sintomi

I segni e sintomi selezionati sono visualizzati in grassetto.

Se ne possono selezionare (deselezionare) altri cliccandoci sopra

**MENÙ AGGIUNGI MALATTIA RARA**

Questa opzione consente di verificare di non avere trascurato una patologia poco probabile (malattia rara) sebbene un sintomo altamente rilevante per quest'ultima fosse stato rilevato.

Tra tutti i sintomi segnalati nella prima fase di indagine vengono visualizzati i primi 10 (valore di default di SiPiuP) sintomi con più alto LR e la patologia per la quale tale valore è massimo.

Se del caso il medico può aggiungere patologie di bassa probabilità nella lista delle malattie sinora già considerate e visualizzarle nel menù successivo.

**Menù Aggiungi Malattia Rara (procedura).**

opzioni possibili:

**Configurazione**

*entra nel menù CPDA*

**Ritorna**

*torna nel menù Home*

l'output visualizza una lista di SiPiuP (10 per default) segni e sintomi ad ognuno di questi è associata la patologia con il massimo LR (interrogando il **DB**) segni e sintomi e patologia già selezionate sono visualizzate in grassetto. Se ne possono selezionare (deselezionare) altre cliccandoci sopra.

**MENÙ SIT**

Menù per la Selezione Interattiva semiautomatica dei Test ad alto effetto diagnostico. Ad ogni patologia della lista selezionata ( LPS: Lista Patologie Selezionate) corrisponde un set specifico di test diagnostici.

Di tutti questi test si considerano solo quelli per cui la probabilità pre-test è minore od uguale al 70% (valore di default della variabile SoSePa\_sup). I test così individuati

compongono il cosiddetto insieme TAMS (test associati alle malattie selezionate).

Dall'insieme TAMS di tutti i test rilevanti vengono dapprima selezionati quelli che in assoluto (cioè per tutte le patologie selezionate) hanno un LR sopra una soglia di *efficacia* SoEf (2 di default) e, successivamente, di questi ultimi, vengono selezionati e visualizzati quelli che, per almeno una delle patologie selezionate superano il valore della probabilità post test del 70% (valore di default della variabile SoPPT e comunque minore od uguale a SoSePa\_sup).

Il medico potrà ulteriormente selezionare/deselezionare i test di TAMS come lo ritenga più opportuno.

**Menù SIT (procedura).**

opzioni possibili:

**Ritorna**

torna al menù precedente

**Configura percorsi diagnostici appropriati**

entra nel menù CPDA

l'output visualizza la lista dei test diagnostici appartenenti all'insieme TAMS con LR > SoEf e probabilità post test > SoPPT.

Si possono selezionare/deselezionare i test cliccandoci sopra.

I test selezionati sono visualizzati in grassetto.

**MENÙ SMT**

Nel menù Selezione Manuale dei Test per la patologia selezionata nel menù precedente vengono visualizzati tutti i test pertinenti ordinati secondo il rapporto di probabilità relativo in modo tale che il medico può selezionare a suo piacimento uno o più test diagnostici. Può quindi accadere che anche un test che non superi la soglia di efficacia possa essere scelto e venire così ad essere inserito nell'insieme TAMS.

**Menù SMT (procedura).**

opzioni possibili:

**Ritorna**

includi i test selezionati nell'insieme TAMS (se non già presenti)

torna al menù precedente

**Configura percorsi diagnostici appropriati**

entra nel menù CPDA

l'output visualizza la lista dei test diagnostici.

Si possono selezionare/deselezionare i test cliccandoci sopra.

I test selezionati sono visualizzati in grassetto.

## **MENÙ CPDA**

Nel menù Configurazione Percorsi Diagnostici Appropriati è consentito visualizzare ed eventualmente modificare il valore delle variabili di sistema.

### **Menù CPDA (procedura).**

opzioni possibili:

#### **NuMaxSint**

20 di default;

numero massimo di segni e sintomi selezionabili per patologia

viene utilizzata nel menù *Lista Sintomi*

#### **NuPaLis**

10 di default; Numero delle Patologie nella Lista.

numero massimo di patologie elencate nel menù

*home* dalla procedura **LP**

#### **NuSePa**

5 di default; Numero massimo della Selezione delle Patologie.

Numero di patologie più probabili considerate nel

menù *Selezione Sintomi Rilevanti*

#### **SiPr**

10 di default; Sintomi Probabili.

Numero massimo di segni e sintomi più probabili per patologia che sono stati ignorati; viene utilizzata nel

menù *Selezione Sintomi Rilevanti*

#### **SiPiuP**

10 di default; Sintomi Più Probabili.

Numero massimo delle patologie per ciascuna delle quali i segni e sintomi selezionati presentano il massimo valore di LR.

Viene usata nel menù *Aggiungi Malattia Rara*

#### **SoSePa\_inf**

40% di default; Soglia Inferiore Selezione Patologia; soglia della

probabilità di una patologia al di sopra della quale la patologia stessa è inclusa nella lista LPP di quelle cioè selezionate automaticamente nel menù *Selezione delle patologie per la diagnosi differenziale*.

È il limite inferiore dell'intervallo di incertezza diagnostica che è dato dalla differenza SoSePa\_sup - SoSePa\_inf.

**SoSePa\_sup**

70% di default; Soglia superiore Selezione Patologia. Limite superiore dell'intervallo dell'incertezza diagnostica che è dato dalla differenza SoSePa\_sup – SoSePa\_inf.

Solo i test diagnostici associati alle patologie le cui probabilità pre-test cadono in questo intervallo vengono presi in considerazione nel menù *Selezione dei Test diagnostici*

**SoEf**

2 di default; Soglia di Efficacia

valore del LR di un test diagnostico al di sopra del quale un test diagnostico può essere selezionato dal medico per elaborare un percorso diagnostico. Viene utilizzata nel menù *SIT*

**SoPPT**

70% di default; Soglia Probabilità Post Test.

Soglia oltre la quale un test può essere incluso dal medico nella lista dei percorsi diagnostici. Deve essere SoPPT <= SoSePa\_sup.

Viene usata nel menù *SIT*

**Te: tempo di attesa per accedere al test**

si seleziona/deseleziona l'utilizzo di tale variabile per il calcolo dell'appropriatezza

**C: indice relativo al costo del test**

si seleziona/deseleziona l'utilizzo di tale variabile per il calcolo dell'appropriatezza

**R: indice relativo al rischio del test**

si seleziona/deseleziona l'utilizzo di tale variabile per il calcolo dell'appropriatezza

**To: indice relativo alla tollerabilità al test**

si seleziona/deseleziona l'utilizzo di tale variabile per il calcolo dell'appropriatezza

### **Visualizza valori appropriatezza**

si seleziona/deseleziona la possibilità di visualizzare esplicitamente i valori di appropriatezza nell'output del menù Selezione dei Test Diagnostici

### **Ritorna**

ritorna nel menù precedente nell'output viene visualizzato il valore aggiornato di ogni variabile.

### **Procedura LP - descrizione**

La procedura LP costruisce la Lista delle Patologie più Probabili (LPP) utilizzando un modello NAIVE BAYES (ad esempio descritto in Waghlikar et al. Su citato).

Si inizia dalla conoscenza della prevalenza, espressa in termini di odds, della patologia cioè dal dato epidemiologico.

Con il termine inglese “odds” in statistica si intende il rapporto tra la probabilità “p” di un evento e la probabilità che tale evento non accada (cioè la probabilità (1-p) dell'evento complementare). Odds\_prevalenza indica il rapporto tra la probabilità che il paziente abbia una patologia in base a considerazioni demografiche e legate all'età (prevalenza) e la probabilità che non abbia quella patologia (1-prevalenza). Definendo con P(M) la prevalenza della patologia M si ha che  $\text{odds\_prevalenza} = P(M) / (1-P(M))$ .

Ad ogni sintomo, segno, test effettuato, dato anamnestico è associato un rapporto di probabilità (likelihood ratio) LR dato da:

$$LR = P(S|M) / P(S|\sim M)$$

e cioè dalla probabilità condizionata di riscontrare il sintomo (S) in un paziente malato (M) diviso la probabilità condizionata di avere lo stesso sintomo in un paziente che non ha quella specifica malattia ( $\sim M$ ). Il LRg globale di un insieme di segni, sintomi, ecc è dato dal prodotto dei singoli LR.

Si può ora stimare la probabilità che un paziente con quei sintomi, segni ecc abbia la malattia M prima di eseguire ulteriori test diagnostici (pre-test); in termini di odds si scrive:

$$\text{odds\_pre\_test} = LRg * \text{odds\_prevalenza}$$

e questo, in generale, si può di nuovo esprimere in termini di probabilità P sapendo che

$P = \text{odds} / (1 + \text{odds})$ .

La formula del calcolo di odds\_pre\_test sopra descritta è di semplice ed immediata implementazione e, come già detto, da quest'ultima si calcola la probabilità P\_pre\_test. Alternativamente, seppur sia leggermente più complicato, si può calcolare direttamente la probabilità pre-test per una successione di n sintomi in maniera iterativa o ricorsiva. In questo caso si utilizza la relazione:

$$P_i = (LR * P_{i-1}) / (LR * P_{i-1} + K_{i-1}) \quad \text{dove } K_{i-1} = (1 - P_{i-1})$$

$P_i$  è la probabilità pre-test dei primi i sintomi; dove i va da 1 a tutti gli n sintomi e  $P_0$  è la prevalenza della patologia. Ovviamente  $P_{\text{pre\_test}} = P_n$

Seppure tale implementazione sembri più complessa, essa, però, consente di implementare semplici varianti al modello rigorosamente bayesiano. Per esempio si possono adottare diverse formulazioni di  $K_{i-1}$  come le due seguenti:

$$K_{i-1} = (1 - P_0) \quad \text{oppure} \quad K_{i-1} = (1 - P_{i-1})^{LR}$$

L'idea di base di queste due ultime formulazioni, è quella di abbassare le probabilità pre-test ricavate dalla formula di Bayes che considera tutte le variabili del sistema indipendenti tra di loro.

### **Procedura LP**

for (ogni patologia della lista "Malattie" dalla **BDC**)

$P_0 =$  prevalenza della patologia

$ODDS = P_0 / (1 - P_0)$  /\* prevalenza espressa in termini di odds \*/

$LRg = 1$  /\*likelihood ratio globale per tutti i dati associati al paziente\*/

for (ogni sintomo, segno, dato anamnestico, test effettuato)

leggi il corrispondente LR (likelihood ratio)

$LRg = LRg * LR$

$ODDSpre = LRg * ODDS$

$PpreTest = ODDSpre / (ODDSpre + 1)$  /\* PpreTest è la probabilità pre-test \*/

memorizza la lista LPP di ogni patologia con la rispettiva PpreTest

end

### **Procedura OPDA - descrizione**

La procedura elabora l'output del percorso diagnostico appropriato utilizzando lo stesso modello NAIVE BAYES e le stesse formule riportate nella procedura LP.

Sia LR il rapporto di probabilità (likelihood ratio) associato ad un generico test

diagnostico, la probabilità stimata (espressa in termini di odds) della patologia M in caso di esito positivo del test (post-test) è data da:

$$\text{odds\_post\_test} = \text{LR} * \text{odds\_pre\_test}$$

e questo si può di nuovo esprimere in termini di probabilità P sapendo che

$$P = \text{odds} / (1 + \text{odds}).$$

Si può valutare questa probabilità per ogni singolo test e per l'insieme di tutti i test.

La procedura associa, inoltre, ad ogni test il rispettivo indice di appropriatezza IA definito come:

$$IA = T_o / (C * T_e * R)$$

dove  $T_o$  è l'indice relativo alla tollerabilità al test;  $C$  è l'indice relativo al costo del test;  $T_e$  è l'indice relativo al tempo di attesa per accedere al test;  $R$  è l'indice relativo al massimo tra il rischio intrinseco e il rischio relativo del test.

Maggiore è l'appropriatezza maggiore è il valore dell'indice IA.

L'indice di appropriatezza globale ( $IA_G$ ) di tutti i test diagnostici selezionati, viene calcolato come:

$$IA_G = T_{oG} / (C_G * T_{eG} * R_G)$$

dove  $T_{oG}$  uguale al minimo tra tutti gli indici di tollerabilità;  $C_G$  uguale alla somma dei singoli costi per ciascun test;  $T_{eG}$  uguale al massimo fra tutti i tempi di attesa;  $R_G$  uguale al massimo fra tutti gli indici di rischio.

Il medico può decidere di definire un percorso alternativo (selezionando un insieme diverso di test), il sistema mantiene in memoria il percorso precedente con i relativi indici di appropriatezza in modo tale che il medico possa confrontare diversi percorsi.

Per il calcolo di  $\text{odds\_post\_test}$  si possono ripetere le stesse considerazioni descritte a commento della procedura LP riguardo al calcolo della probabilità  $P_{\text{pre\_test}}$  per quanto riguarda l'indipendenza statistica delle variabili associate ai segni, sintomi e test. E valgono anche in questo caso le varianti al calcolo rigorosamente bayesiano lì descritte.

### **Procedura OPDA**

for (ogni patologia della lista LPS dalla **BDC**)

$$P_0 = P_{\text{preTest}}$$

/\* probabilità pre test della patologia calcolata in precedenza \*/

$$\text{ODDS} = P_0 / (1 - P_0)$$

$$\text{LR}_g = 1 \quad /* \text{likelihood ratio globale per tutti i test} */$$

for (ogni test appartenente all'insieme TAMS )

leggi il corrispondente LR (likelihood ratio)

$$LR_g = LR_g * LR$$

leggi  $T_e$  /\* indice relativo al tempo di attesa per accedere al test \*/

leggi  $C$  /\* indice relativo al costo del test \*/

leggi  $R_i$  /\* indice relativo al rischio intrinseco del test \*/

leggi  $R_r$  /\* indice relativo al rischio relativo del test \*/

determina  $R$  /\* come il massimo tra  $R_i$  e  $R_r$  del test \*/

leggi  $T_o$  /\* indice relativo alla tollerabilità al test \*/

calcola  $C_G$  come somma di tutti i  $C$

determina  $T_{eG}$  come il massimo dei singoli  $T_e$

determina  $R_G$  come il massimo dei singoli  $R$

determina  $T_{oG}$  come il minimo dei singoli  $T_o$

Appropriatezza =  $T_o / (C * T_e * R)$  /\* indice di appropriatezza del test \*/

$IA_G = T_{oG} / (C_G * T_{eG} * R_G)$  /\* appropriatezza globale di tutti i test \*/

$$K = LR * ODDS$$

$P_{postTest} = K / (1 + K)$  /\* probabilità post-test per ogni test \*/

$$ODDS_{postG} = LR_g * ODDS$$

/\* calcola la probabilità post-test globale \*/

$$P_{postTestG} = ODDS_{postG} / (1 + ODDS_{postG})$$

for (ogni patologia della lista LPS dalla BDC)

memorizza e visualizza:

la denominazione della patologia e la rispettiva  $P_{preTest}$

memorizza e visualizza in ordine decrescente di appropriatezza

ogni sintomo con la rispettiva  $P_{postTest}$

e (opzionalmente) l'indice di appropriatezza

la probabilità post-test globale:  $P_{postTestG}$

l'appropriatezza globale di tutti i test (se l'opzione è attiva)

end

## Glossario

BDC: Base Di Conoscenza

C: costo del test diagnostico (variabile)

CPDA: Configurazione Percorso Diagnostico Appropriato (menu)

DB: database

DP: dati paziente

LMR: lista delle Malattie Rare

LP: procedura di calcolo delle probabilità delle ipotesi diagnostiche (procedura)

LPA: Lista Patologie Aggiunte manualmente

LPP: Lista di Patologie più Probabili

LPS: Lista delle Patologie Selezionate

LR: Likelihood Ratio

NuMaxSint: Numero Massimo di segni e Sintomi selezionabili per patologia (variabile)

NuPaLis: Numero di Patologie nella Lista (variabile)

NuSePa: Numero Massimo della Selezione di Patologie (variabile)

OPDA: Output Percorso Diagnostico Appropriato

OPDA: procedura di elaborazione dell'output del PDA (procedura)

R: Rischio associato al test (variabile)

Ri: Rischio intrinseco del test (variabile)

Rr: Rischio relativo del test (variabile)

SiPiuP: Sintomi più Probabili (variabile)

SiPr: Sintomi Probabili (variabile)

SIT: Selezione Interattiva dei Test diagnostici (Menu)

SMT: Selezione Manuale dei Test diagnostici (Menu)

SoEf: Soglia di Efficacia (variabile)

SoPPT: Soglia di Probabilità Post-Test (variabile)

SoSePa \_inf: Soglia inferiore per la Selezione delle Patologie (variabile)

SoSePa \_sup: Soglia superiore per la Selezione delle Patologie (variabile)

TAMS: l'insieme dei test associati alle malattie selezionate

Te: tempo di accesso al test diagnostico (variabile)

To: indice di tollerabilità del test (variabile)

La presente invenzione può essere vantaggiosamente realizzata tramite un programma per computer che comprende mezzi di codifica per la realizzazione di uno o più passi del metodo, quando questo programma è eseguito su di un computer. Pertanto si intende che l'ambito di protezione si estende a detto programma per computer ed inoltre a mezzi

leggibili da computer che comprendono un messaggio registrato, detti mezzi leggibili da computer comprendendo mezzi di codifica di programma per la realizzazione di uno o più passi del metodo, quando detto programma è eseguito su di un computer.

Sono possibili varianti realizzative all'esempio non limitativo descritto, senza per altro uscire dall'ambito di protezione della presente invenzione, comprendendo tutte le realizzazioni equivalenti per un tecnico del ramo.

Gli elementi e le caratteristiche illustrate nelle diverse forme di realizzazione preferite possono essere combinati tra loro senza peraltro uscire dall'ambito di protezione della presente invenzione.

Dalla descrizione sopra riportata il tecnico del ramo è in grado di realizzare l'oggetto dell'invenzione senza introdurre ulteriori dettagli costruttivi.

## RIVENDICAZIONI

1. Apparato di diagnosi differenziale adattato per applicazione in medicina per determinare una sequenza ottimale di test diagnostici atta a identificare una patologia adottando criteri di appropriatezza diagnostica, comprendente:

- una prima base dati aggiornabile contenente dati dei pazienti;
- una seconda base dati di tipo relazionale aggiornabile contenente dati identificativi di patologie, sintomi, segni clinici, dati identificativi di test diagnostici, e dati relativi ai parametri di appropriatezza di detti test diagnostici per definire una lista di ipotesi diagnostiche (patologie) ;
- mezzi atti a determinare detta sequenza ottimale di test diagnostici atta a identificare una patologia, detti mezzi comprendendo un motore di calcolo di tipo inferenziale, che sulla base di dati contenuti in dette prima e seconda base dati, determinano per ogni ipotesi diagnostica (patologia) detta sequenza ottimale di test diagnostici con associati relativi indici di appropriatezza e probabilità che un paziente sia affetto da quella patologia.

2. Apparato di diagnosi come nella rivendicazione 1, in cui detti mezzi atti a determinare detta sequenza ottimale di test diagnostici sono atti a, tramite detto motore di calcolo di tipo inferenziale:

- determinare, a partire da detti dati contenuti in dette prima e seconda base dati, un indice di probabilità pre-test delle possibili ipotesi diagnostiche (patologie)
- determinare, a partire da detti dati contenuti in dette prima e seconda base dati e da detto indice di probabilità pre-test, un indice di probabilità post-test delle possibili ipotesi diagnostiche (patologie) condizionata all'esecuzione di detta sequenza ottimale di test diagnostici;
- associare un indice di appropriatezza ad ogni test di detta sequenza ottimale;
- associare un indice di appropriatezza a detta sequenza ottimale di test diagnostici.

3. Apparato di diagnosi come nella rivendicazione 2, in cui detti mezzi atti a determinare detta sequenza ottimale di test diagnostici sono atti a determinare detta probabilità pre-test  $Pre(M | \{DP\})$  e detta probabilità post-test  $Post(M | \{DP\} \cup \{T\})$  in base a detto calcolo di tipo inferenziale di rispettive probabilità pre-test e post test di riscontrare detta patologia  $M$  condizionata rispettivamente all'insieme dei dati identificativi del paziente contenuti in detta prima base dati  $\{DP\}$  e all'unione di detti dati identificativi del paziente  $\{DP\}$  con l'insieme di detti test  $\{T\}$  di detta sequenza ottimale di test

diagnostici.

4. Apparato di diagnosi come nella rivendicazione 3, in cui detti mezzi atti a determinare detta sequenza ottimale di test diagnostici comprendono mezzi atti a determinare detta probabilità pre-test  $Pre(M | \{DP\})$  in modo tale che:

- associare ad ogni sintomo un rapporto di probabilità LR dato da:

$$LR = P(S|M) / P(S| \sim M)$$

cioè dalla probabilità condizionata  $P(S|M)$  di riscontrare il sintomo (S) in un paziente affetto da una patologia (M) diviso la probabilità condizionata  $P(S| \sim M)$  di avere lo stesso sintomo in un paziente non affetto da detta patologia ( $\sim M$ );

- calcolare un rapporto di probabilità globale LRg dato dal prodotto dei rapporti di probabilità LR di ogni patologia;

- calcolare detta probabilità pre-test  $Pre(M | \{DP\})$  relativa dapprima in termini di odds:

$$odds\_pre\_test = LRg * odds\_prevalenza$$

dove  $odds\_prevalenza = P(M) / (1 - P(M))$  in cui  $P(M)$  è la prevalenza della patologia M.

da cui detta probabilità pre-test  $Pre(M | \{DP\})$  è data da:

$$Pre(M | \{DP\}) = odds / (1 + odds).$$

5. Apparato di diagnosi come nella rivendicazione 4, in cui detti mezzi atti a determinare detta sequenza ottimale di test diagnostici comprendono mezzi atti a determinare detta probabilità post-test  $Post(M | \{DP\} \cup \{T\})$  in modo tale che:

$$odds\_post\_test = LR * odds\_pre\_test$$

da cui detta probabilità post-test  $Post(M | \{DP\} \cup \{T\})$  è data da:

$$Post(M | \{DP\} \cup \{T\}) = odds\_post\_test / (1 + odds\_post\_test).$$

6. Apparato di diagnosi come nella rivendicazione 3, in cui detti mezzi atti a determinare detta sequenza ottimale di test diagnostici comprendono mezzi atti a determinare detta probabilità pre-test  $Pre(M | \{DP\})$  per una successione di sintomi in maniera iterativa o ricorsiva, utilizzando la relazione:

$$P_i = (LR * P_{i-1}) / (LR * P_{i-1} + K_{i-1}) \text{ dove } K_{i-1} = (1 - P_{i-1})$$

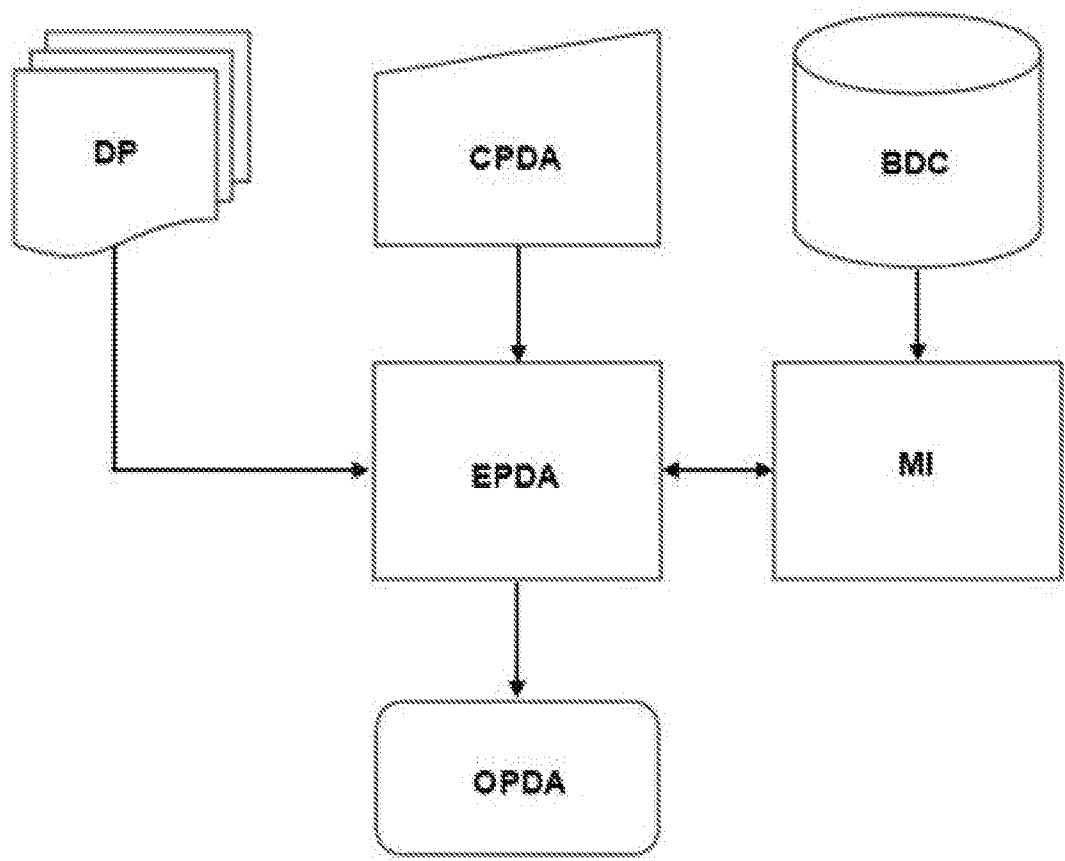
$P_i$  è la probabilità pre-test dei primi i sintomi; dove i va da 1 a tutti gli n sintomi e  $P_0$  è la prevalenza della patologia;

- quindi determinare detta probabilità pre-test  $Pre(M | \{DP\}) = P_n$

7. Apparato di diagnosi come nella rivendicazione 1 in cui detti mezzi atti a determinare detta sequenza ottimale di test diagnostici comprendono mezzi atti a:

- determinare detto indice di appropriatezza IA di un singolo test diagnostico definito come:  $IA = T_o / (C * T_e * R)$ , dove  $T_e$  è l'indice relativo al tempo di attesa per accedere al test ;  $C$  è l'indice relativo al costo del test;  $R$  è l'indice relativo al massimo tra il rischio intrinseco e il rischio relativo del test;  $T_o$  è l'indice relativo alla tollerabilità al test;
- determinare detto indice di appropriatezza globale di una sequenza di test diagnostici  $IA_G = T_{oG} / (C_G * T_{eG} * R_G)$  dove  $T_{oG}$  è pari al minimo tra tutti gli indici di tollerabilità tra tutti i test della sequenza,  $C_G$  è pari alla somma dei singoli costi di ciascun test della sequenza,  $T_{eG}$  è pari al massimo fra tutti tempi di attesa,  $R_G$  è pari al massimo fra tutti gli indici di rischio della sequenza.

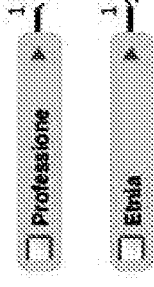
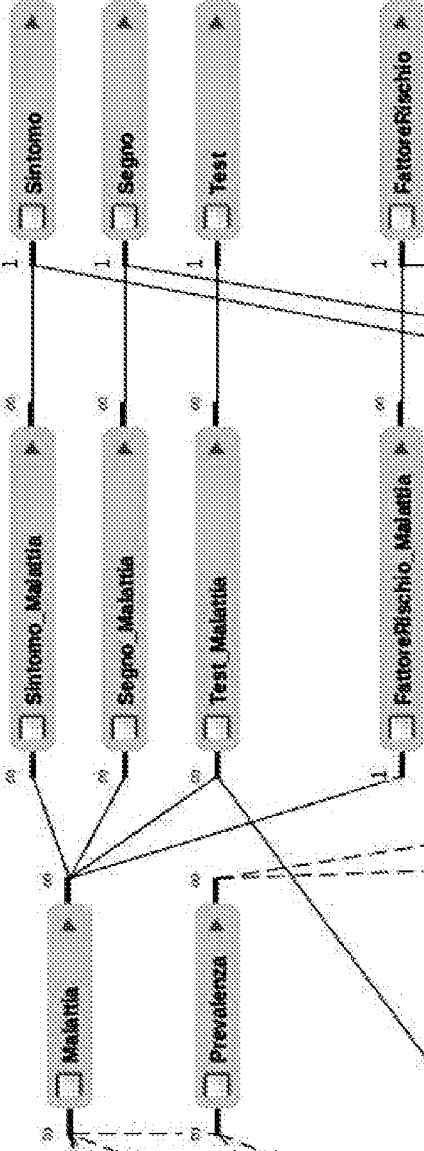
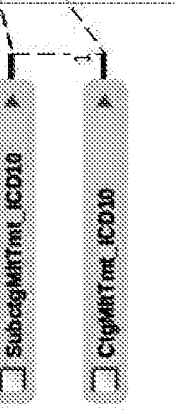
8. Apparato di diagnosi come nella rivendicazione 1, comprendente moduli software implementati in dispositivi di memorizzazione leggibili ed eseguibili da un computer corredato da dispositivi di input/output, implementati su cluster di calcolo (fisico o virtuale) ed eseguibili da remoto ed implementati come App per dispositivi mobili (Tablet, smartphone, PC portatili).



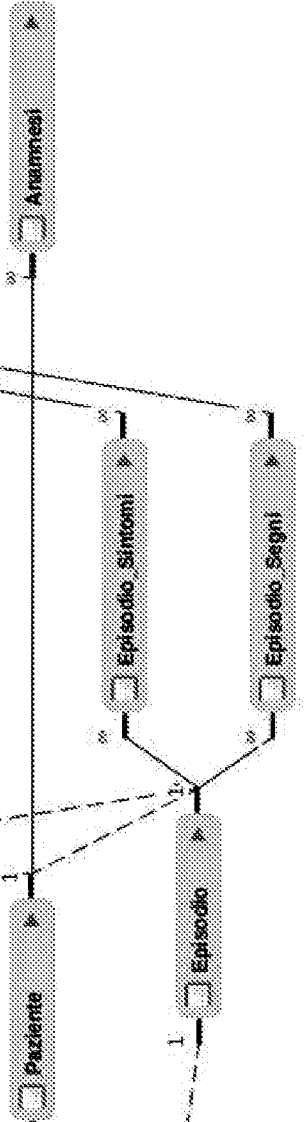
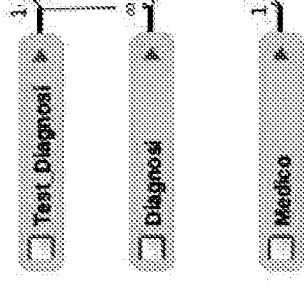
**FIG. 1**

**BDC**

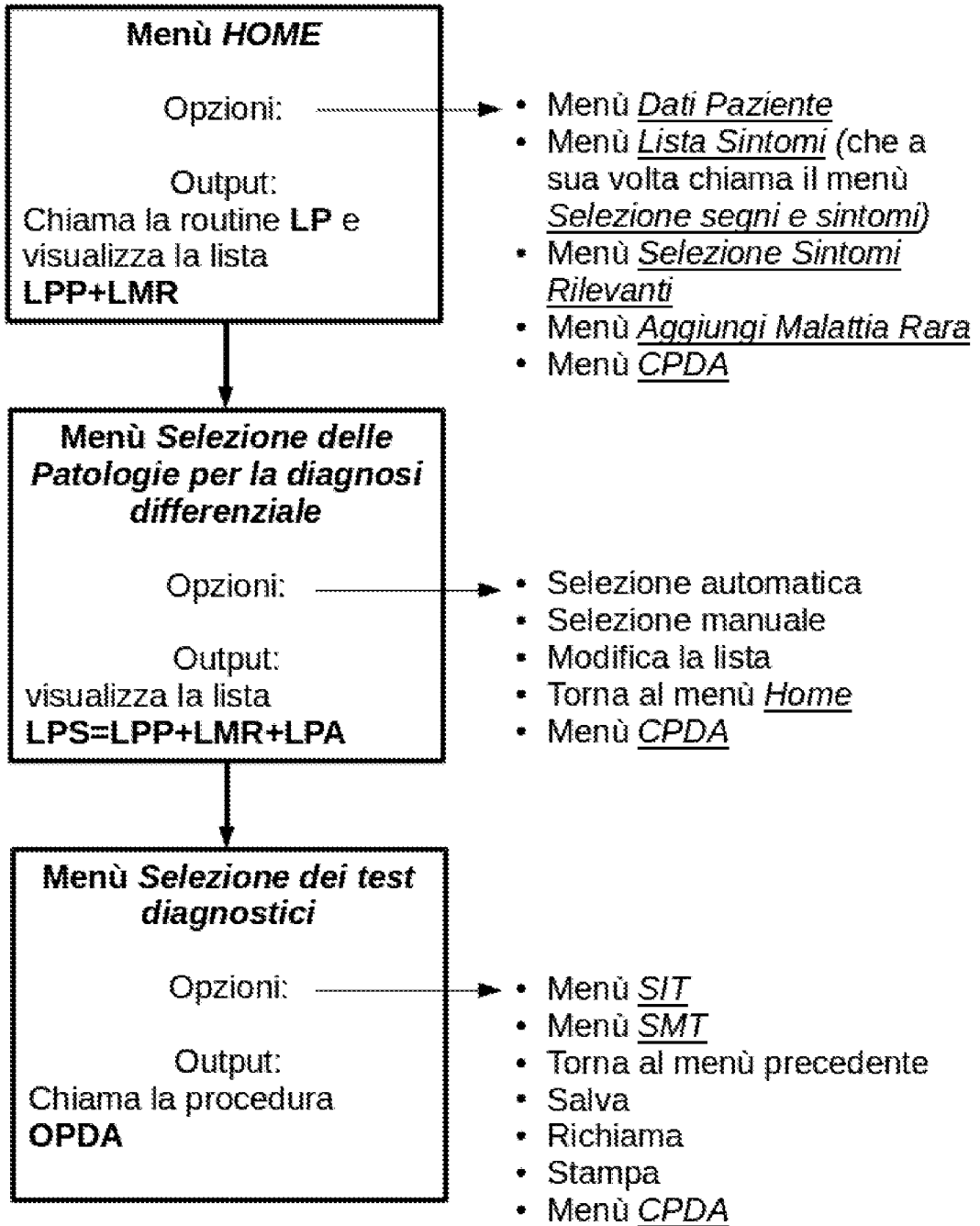
**ICD10**



**PDA**



**FIG. 2**



**FIG. 3**