

Report del progetto
AMD KIDNEY DIGITAL TWIN:

focus su
2 Modelli di Previsione
Stima eGFR dopo 3 anni in caso di
utilizzo o non utilizzo di SGLT2i



SOMMARIO

EPIDEMIOLOGIA DELLA MALATTIA RENALE CRONICA NEL DIABETE	3
IL TRATTAMENTO FARMACOLOGICO NEL CONTROLLO DELLA PROGRESSIONE DELLA MALATTIA RENALE CRONICA NEL DIABETE	3
GLI STUDI EPIDEMIOLOGICI ITALIANI	4
PREVENZIONE E DIAGNOSI PRECOCE DELLA MRC NEL DIABETE	5
DIGITAL TWIN: UNO STRUMENTO PER LA PREVENZIONE PERSONALIZZATA NEL DIABETE	5
IL PROGETTO PROPOSTO: CREARE I PRESUPPOSTI PER LA CREAZIONE DI UN ‘AMD KIDNEY DIGITAL TWIN’ A SUPPORTO DEL DIABETOLOGO	6
METODOLOGIA.....	7
FONTE DATI.....	7
IL METODO UTILIZZATO NELL’ANALISI: LLM (LOGIC LEARNING MACHINE), UNO STRUMENTO BASATO SUL MACHINE LEARNING EXPLAINABLE A REGOLE ESPLICITE.....	8
LA AI TRASPARENTE SU CUI SI BASA IL PROGETTO: GARANZIA DI AUGMENTED INTELLIGENCE, INFORMED DECISIONS E COMPLIANCE CON GDPR, AI ACT E LEGGE ITALIANA 132/2025.....	8
LA STIMA DEL VALORE DI EGFR DOPO 3 ANNI, NELLE 2 SITUAZIONI DI UTILIZZO/NON UTILIZZO DI SGLT2.....	9
CRITERI DI SELEZIONE E NUMEROSITÀ DEL CAMPIONE.....	9
MODELLI DI REGRESSIONE STIME EGFR “SGLT2-SI” E “SGLT2-NO”: RISULTATI	12
L’APPLICAZIONE AMD KDT (KIDNEY DIGITAL TWIN).....	12
SCOPO DELL’APPLICAZIONE AMD KDT (KIDNEY DIGITAL TWIN)	12
FUNZIONI PRINCIPALI	12
UTENTI E CONTESTO D’USO	13
CONOSCENZA SU CUI SI BASA L’APPLICAZIONE	13
AUTORIZZAZIONI, AMMINISTRAZIONE E REQUISITI DI PRIVACY.....	13
FUNZIONALITÀ PRINCIPALI LATO UTENTE	14
BIBLIOGRAFIA.....	12

Epidemiologia della malattia renale cronica nel diabete

La malattia renale cronica (CKD) rappresenta una delle complicanze più frequenti e gravi del diabete mellito. La prevalenza di CKD tra le persone con diabete è stimata tra il 20% e il 40%, con una tendenza all'aumento proporzionale all'incremento globale della prevalenza del diabete stesso. Si stima che circa il 40% delle persone con diabete svilupperà CKD nel corso della vita. Il diabete è la causa principale di insufficienza renale terminale a livello mondiale e rappresenta circa la metà dei nuovi casi di insufficienza renale negli Stati Uniti. La CKD nel diabete si manifesta tipicamente dopo 5-15 anni dalla diagnosi di diabete tipo 1, mentre può essere già presente al momento della diagnosi di diabete tipo 2. La presenza di CKD nei pazienti diabetici aumenta significativamente il rischio di eventi cardiovascolari, insufficienza cardiaca e mortalità[1].

I principali fattori di rischio modificabili per lo sviluppo e la progressione della malattia renale cronica nei pazienti con diabete sono:

1. Controllo glicemico subottimale: L'iperglicemia persistente è il fattore causale più importante; il mantenimento di un target di HbA1c inferiore al 7% riduce il rischio di progressione della nefropatia diabetica[1].
2. Ipertensione arteriosa: La pressione arteriosa elevata accelera la progressione della CKD; il controllo pressorio, preferibilmente con ACE-inibitori o ARB, è raccomandato soprattutto in presenza di albuminuria[1].
3. Albuminuria: la presenza di albuminuria è sia marker che fattore di rischio; la riduzione dell'albuminuria tramite RAS-inibitori, SGLT2-inibitori e antagonisti dei recettori mineralcorticoidi (come finerenone) rallenta la progressione[1].
4. Dislipidemia: l'ipercolesterolemia e l'ipertrigliceridemia sono associati a peggioramento della funzione renale; l'uso di statine è raccomandato per tutti i pazienti con diabete e CKD[1].
5. Obesità: l'eccesso ponderale favorisce l'insulino-resistenza e l'ipertensione; la riduzione del peso corporeo migliora il controllo pressorio e metabolico[2].
6. Dieta ricca di sodio e proteine animali: l'eccesso di sodio aumenta la pressione arteriosa e l'albuminuria; una dieta con restrizione di sodio e preferenza per proteine vegetali può essere benefica[2].
7. Fumo di sigaretta: il fumo è associato a maggiore incidenza e progressione di CKD; la cessazione è fortemente raccomandata[2].
8. Sedentarietà: l'attività fisica regolare è associata a minore progressione della CKD e riduzione del rischio cardiovascolare[2].

La gestione multifattoriale prevede l'ottimizzazione di tutti questi fattori tramite modifiche dello stile di vita e terapie farmacologiche mirate[1].

Il trattamento farmacologico nel controllo della progressione della malattia renale cronica nel diabete

Gli ACE-inibitori e gli ARB rimangono la terapia di base, come dimostrato nei trial IDNT (irbesartan) [3] RENAAL (losartan) [4] con effetti protettivi sulla progressione della nefropatia diabetica.

L'uso combinato di SGLT2-inibitori, GLP-1 agonisti e ns-MRA, oltre al controllo pressorio e glicemico, rappresentano l'approccio più efficace secondo le evidenze più recenti[1],[5] [6]

Gli studi più recenti evidenziano che gli SGLT2-inibitori sono il trattamento farmacologico con la più solida efficacia nella prevenzione della progressione della malattia renale cronica nei pazienti diabetici, con benefici che includono la riduzione del rischio di insufficienza renale terminale, morte per cause renali e cardiovascolari, e ospedalizzazione per scompenso cardiaco. Questi effetti sono stati confermati da numerosi trial randomizzati e metanalisi, e la raccomandazione di prima linea è sostenuta sia dalle linee guida Kidney Disease: Improving Global Outcomes (KDIGO) che

dall'American Diabetes Association, che suggeriscono l'uso di SGLT2-inibitori in pazienti con diabete tipo 2, CKD e eGFR ≥ 20 ml/min/1,73m² [1][5].

Gli SGLT2-inibitori sono stati valutati in studi di grande rilievo come CREDENCE (canagliflozin) [7] DAPA-CKD (dapagliflozin) [8] ed EMPA-KIDNEY (empagliflozin) [9], [10] dimostrando una significativa riduzione del rischio di progressione verso l'insufficienza renale terminale, declino sostenuto dell'eGFR e morte cardiovascolare. Questi benefici si osservano anche in pazienti con CKD avanzata.

Gli agonisti del recettore GLP-1 offrono benefici aggiuntivi, in particolare nella riduzione degli eventi cardiovascolari e della progressione dell'albuminuria; tuttavia, le evidenze relative alla loro efficacia sulla progressione verso l'insufficienza renale terminale sono più limitate rispetto a quelle disponibili per gli SGLT2-inibitori.

Gli agonisti del recettore GLP-1, come semaglutide e liraglutide, hanno mostrato riduzione dell'albuminuria e rallentamento del declino dell'eGFR in studi come FLOW [11], LEADER [12] e SUSTAIN-6 [13], anche se le evidenze sull'efficacia nella progressione verso l'ESKD sono inferiori rispetto agli SGLT2-inibitori. Nei pazienti con CKD sono raccomandati come opzione aggiuntiva o alternativa, in particolare nei pazienti con elevato rischio cardiovascolare, presenza di albuminuria persistente, obesità o scarso controllo glicemico, nonché nei soggetti con controindicazione o intolleranza agli SGLT2-inibitori [14] [1] [5].

Gli antagonisti non steroidei del recettore dei mineralcorticoidi (ns-MRA), come finerenone, sono indicati nei pazienti con albuminuria persistente nonostante terapia massimale con ACEi/ARB, eGFR ≥ 25 ml/min/1,73 m² e potassiemia normale. Offrono una riduzione del rischio di progressione renale e di eventi cardiovascolari, tuttavia il beneficio è generalmente inferiore rispetto agli SGLT2-inibitori e il rischio di iperkaliemia richiede monitoraggio [1] [5] [15].

Gli antagonisti non steroidei del recettore dei mineralcorticoidi (ns-MRA), in particolare finerenone, sono stati valutati nei trial FIDELIO-DKD [16] e FIGARO-DKD [17], mostrando riduzione del rischio di progressione renale e di eventi cardiovascolari in pazienti con diabete e CKD.

In sintesi: gli ACE-inibitori e gli ARB rimangono la terapia di base con effetti protettivi sulla progressione della nefropatia diabetica. L'uso combinato di SGLT2-inibitori, GLP-1 agonisti e ns-MRA, oltre al controllo pressorio e glicemico, rappresentano l'approccio più efficace secondo le evidenze più recenti.

Gli studi epidemiologici italiani

I principali studi epidemiologici pubblicati in Italia sulla malattia renale cronica nei pazienti con diabete includono il RIACE Italian Multicentre Study, che ha analizzato oltre 15.000 soggetti con diabete tipo 2, evidenziando che la forma non albuminurica di insufficienza renale è prevalente, soprattutto nelle donne, e che la CKD si associa a una maggiore variabilità dell'HbA1c e a un profilo di rischio cardiovascolare sfavorevole [18][19].

Analisi longitudinali su dati degli Annali dell'Associazione Medici Diabetologi (AMD) hanno evidenziato un aumento degli anni vissuti con disabilità attribuibili a CKD nei diabetici, con trend in crescita soprattutto tra gli uomini [20]. Gli studi derivati dagli Annali AMD hanno fornito una panoramica dettagliata sull'epidemiologia e la progressione della malattia renale cronica nei pazienti con diabete in Italia.

Nei pazienti con diabete di tipo 2, la forma non albuminurica di danno renale è risultata la più frequente, contribuendo in modo significativo al carico di morbidità da insufficienza renale terminale, sebbene associata a un rischio di progressione inferiore rispetto alla nefropatia con albuminuria. La variabilità intra-individuale di HbA1c, lipidi, uricemia e pressione arteriosa si sono dimostrati un determinante più importante nella progressione della CKD rispetto ai valori medi di ciascun parametro. Fattori come ipertensione, iperuricemia, trigliceridi elevati e HDL basso sono risultati

predittori indipendenti di insorgenza di CKD, mentre il controllo pressorio ottimale riduce l'albuminuria ma non elimina il rischio di progressione, con possibile effetto J-curve sull'eGFR a valori pressori molto bassi [21][22].

In particolare, uno studio pubblicato recentemente [23] ha analizzato la prevalenza e i determinanti clinici del rapido declino del tasso di filtrazione glomerulare stimato (eGFR) in pazienti con diabete di tipo 2 (T2D) di nuova diagnosi. Il declino rapido dell'eGFR, definito come una riduzione superiore a 5 mL/min/1.73 m² all'anno, è un predittore indipendente di malattia renale terminale (ESKD) e mortalità.

Lo studio ha analizzato i dati di 105.163 pazienti con T2D di nuova diagnosi provenienti dal database Annali AMD. I risultati mostrano che il 12,9% dei pazienti ha sperimentato un rapido declino dell'eGFR in un periodo di follow-up di 3 anni. I principali fattori predittivi di questo declino includono età avanzata, sesso femminile, livelli elevati di HbA1c, fumo, eGFR iniziale elevato, albuminuria e retinopatia. Lo studio sottolinea l'importanza di monitorare le variazioni dinamiche dell'eGFR e di includerle nella definizione di malattia renale cronica (CKD), per una gestione più accurata dei pazienti a rischio. Inoltre, evidenzia la necessità di strategie preventive mirate per ridurre il rischio di declino rapido della funzione renale.

Prevenzione e diagnosi precoce della MRC nel diabete

Nel corso degli ultimi decenni, la comunità scientifica internazionale ha promosso un ampio ventaglio di iniziative volte a potenziare la prevenzione e la diagnosi precoce della malattia renale cronica (MRC) in pazienti affetti da diabete mellito. Tra le principali strategie adottate si annoverano l'aggiornamento continuo delle linee guida evidence-based, [1][5] la formulazione di documenti di consenso multidisciplinari [24] e l'integrazione di sistemi informatizzati di cartella clinica elettronica. [25][26][27]

Questi interventi sono finalizzati a ottimizzare il monitoraggio dei biomarcatori di funzione renale, favorire l'individuazione precoce del danno renale incipiente e promuovere un approccio terapeutico tempestivo, personalizzato e fondato sulle migliori evidenze disponibili.

Digital twin: uno strumento per la prevenzione personalizzata nel diabete

Negli ultimi anni, con obiettivi di prevenzione, anticipazione del rischio e personalizzazione delle strategie terapeutiche, si è progressivamente affermato il paradigma del *Health Digital Twin* (HDT). L'HDT costituisce una replica digitale di un'entità fisica — individuo, organo o funzione — in grado di riprodurre il comportamento in differenti condizioni fisiopatologiche [28][29].

Basato su modelli di simulazione avanzati e sull'integrazione di algoritmi di intelligenza artificiale (AI) e *machine learning* (ML), l'HDT elabora e interpreta dati clinici e operativi per generare previsioni a supporto del processo decisionale medico. L'aggiornamento continuo, anche in tempo reale, consente un monitoraggio dinamico dei parametri critici e l'attuazione di interventi precoci finalizzati alla prevenzione di eventi avversi o complicanze.

In ambito sanitario, l'HDT rappresenta uno strumento emergente per la gestione predittiva e personalizzata della malattia [30]. In diabetologia, il suo potenziale risiede nel miglioramento del controllo glicemico e nella riduzione del rischio di complicanze croniche, in particolare di natura renale [31][32].

Il progetto proposto: creare i presupposti per la creazione di un 'AMD Kidney Digital Twin' a supporto del diabetologo

Proponiamo di effettuare un'analisi dati e uno studio di fattibilità basato sull'Intelligenza Artificiale per definire l'architettura di un 'AMD Kidney Digital Twin' (AMD-KDT), per eventuali sviluppi da valutare successivamente. L'AMD-KDT si potrebbe configurare come uno strumento a valenza scientifica, educativa e formativa, pensato per potenziare e rendere più agevoli le decisioni informate del diabetologo, con l'obiettivo clinico di preservare il più a lungo possibile la funzione renale nel paziente con diabete di tipo 2.

Gli aspetti indirizzati in questo progetto si riferiscono a:

1. LA PREVISIONE DEL RISCHIO D'INSORGENZA DI MRC;
2. LA STIMA SUL PROBABILE ANDAMENTO DELL'eGFR DOPO 3 ANNI DALLA VISITA CORRENTE (IN CASO DI INTRODUZIONE/NON INTRODUZIONE DI SGLT2i).

PER QUANTO RIGUARDA L'OBIETTIVO 1 (VALUTAZIONE RISCHIO D'INSORGENZA DI MRC) SONO PREVISTE LE SEGUENTI ATTIVITÀ:

1. Creazione di un modello di Machine Learning di classificazione con obiettivi CONOSCITIVI:

- Individuazione delle caratteristiche correlate a 2 tipologie di pazienti: rischio MRC entro 3 anni SI/NO
- Esplicitazione dell'affidabilità del modello
- Esplicitazione (Al trasparente) delle caratteristiche individuate per i due sottogruppi di pazienti
- Fotografia di su una serie di variabili al baseline (3 anni prima del manifestarsi della MRC) e al manifestarsi della MRC

2. Descrizione degli aspetti architetturali e funzionali di una possibile applicazione di KIDNEY DIGITAL TWIN, CON OBIETTIVI EDUCATIVI E SCIENTIFICO-FORMATIVI, che fornisca:

- Una previsione: responso in real time sul probabile rischio d'insorgenza di MRC entro 3 anni dalla visita corrente per un paziente con DT2
- L'esplicitazione delle motivazioni correlate al responso
- L'eventuale bibliografia a supporto

PER QUANTO RIGUARDA L'OBIETTIVO 2, (STIMA PROBABILE ANDAMENTO DELL'EGFR DOPO 3 ANNI), SONO PREVISTE LE SEGUENTI ATTIVITÀ:

1. Creazione modelli di ML di regressione con obiettivi PREDITTIVI (CONOSCITIVI)

- 2 modelli predittivi che stimano il valore di eGFR (o la sua pendenza-slope) dopo 3 anni nelle 2 casistiche:
 - prescrizione di SGLT2i alla visita → SI
 - prescrizione di SGLT2i alla visita → NO
- Esplicitazione affidabilità del modello
- Esplicitazione delle variabili che influenzano il fenomeno analizzato
- Fotografie con dettagli su una serie di variabili al baseline (prescrizione SGLT2i) e dopo 3 anni

3. Descrizione degli aspetti architetturali e funzionali di una possibile applicazione di KIDNEY DIGITAL TWIN, CON OBIETTIVI EDUCATIVI E SCIENTIFICO-FORMATIVI, che fornisca:

- Stima valore eGFR dopo 3 anni in caso di utilizzo di SGLT2i alla visita → SI

- Stima valore eGFR dopo 3 anni in caso di utilizzo di SGLT2i alla visita → NO

Metodologia

Fonte dati

Il progetto proposto si basa su database e dati di proprietà esclusiva di AMD. Nello specifico, si fa riferimento al DB Annali 2023 che contiene:

- 1.587.873 soggetti con diabete tipo 2, 112.041 con diabete tipo 1, 3.769 LADA, 171.127 diabete gestazionale e 66.238 con altre forme di diabete (periodo 2010-2023)
- Si riportano a titolo esemplificativo le numerosità di alcune fra le variabili raccolte nel DB Annali AMD:
- 11.585.690 valori di peso
- 3.087.942 circonferenze vita
- 9.775.863 valori pressori
- 17.565.628 HbA1c
- 180.482.739 dati di laboratorio
- 91.478.941 dati di farmaci
- 145.787.409 complicanze e comorbidità

NB: laddove, negli ambiti di progetto proposti, si preveda di creare modelli, algoritmi e software, essi rimangono di proprietà di AMD. Al cliente viene consegnata la reportistica che illustra metodologia e risultati prodotti.

Le tipologie di dati contenuti nel DB Annali e utilizzati nelle analisi previste per questo progetto sono sintetizzati alla tabella che segue:

PARAMETRI ANTROPOMETRICI	PARAMETRI ORGANIZZATIVI
Età	Q-score
Sesso	Anni di osservazione presso il centro di diabetologia
Peso	Durata della malattia
Altezza	TRATTAMENTI
Body mass index (BMI)	Terapie per il diabete (tipo e associazioni)
Pressione Arteriosa (Sistolica, Diastolica)	Terapie ipolipemizzanti (tipo e associazioni)
<i>(e relative variabili derivate o indici calcolati in base ai parametri sopra elencati)</i>	Terapie antipertensive (presenza, tipo e associazioni)
PARAMETRI BIOCHIMICI	Terapie per la tiroide
HbA1c	COMPLICANZE E COMORBIDITÀ
Glicemia	Nefropatia
Trigliceridi	Fibrillazione atriale
Colesterolo totale	Scompenso cardiaco
HDL	Ictus
LDL	Complicanze CVD
Creatinina	Vasculopatia
eGFR	Arteropatia periferica
Micro–Macro/albuminuria	Neuropatia
Acido urico sierico	Complicanze piede
GOT	Complicanze occhio
GPT	Epatopatia

Il metodo utilizzato nell'analisi: LLM (Logic Learning Machine), uno strumento basato sul Machine Learning Explainable a regole esplicite

Il Machine Learning (ML) è in grado di analizzare i dati senza formulare ipotesi a priori, con la possibilità di prevedere valori di output su nuovi dati, cioè su dati non utilizzati durante l'addestramento.

Una particolare tecnica di ML, la LLM (Logic Learning Machine), si basa su algoritmi nativamente trasparenti (clear box o glass box), costruendo modelli composti da insiemi di regole intelligibili che permettono di estrarre conoscenze sulle variabili analizzate e sulle loro relazioni, rispetto ad un outcome definito (es. previsione dell'evoluzione della patologia, previsione di complicanze, etc).

In letteratura esistono due paradigmi principali per la generazione di regole: i decision tree, che adottano un approccio del tipo divide-and-conquer per generare il modello finale, e i metodi basati sulla ricostruzione di funzioni booleane, che costruiscono l'insieme delle regole seguendo una logica aggregativa.

LLM è un algoritmo proprietario che implementa il modello di switching neural network per risolvere problemi di classificazione, producendo insiemi di regole intelligibili espresse nella forma: "IF premessa..., THEN conseguenza...". La premessa contiene una o più condizioni sulle variabili in input, mentre la conseguenza rappresenta il valore di output o l'outcome definito come obiettivo (ad es. 'rischio SI/NO') [33][34][35][36].

La tecnica di generazione delle regole su cui si basa la LLM identifica un insieme di variabili particolarmente rilevanti, associate a uno specifico esito (ovvero, che influenzano l'outcome), e fa emergere, in modo esplicito e comprensibile, le condizioni relative a tale esito, indicando anche le soglie importanti per ciascuna variabile in input, ovvero, i cosiddetti cut-off ai quali si inverte la probabilità di appartenere ad una certa classe di output o un'altra (es. 'rischio SI/NO'). L'approccio "clear box" adottato dalla LLM consente un'autentica explainable AI, in grado di fornire risultati trasparenti, interpretabili e affidabili, così come generati dall'algoritmo di ML, senza necessità di applicare strumenti di 'decriptazione' degli algoritmi, come ad esempio il metodo SHAP, basato sui valori di Shapley.

La AI trasparente su cui si basa il progetto: garanzia di Augmented Intelligence, Informed Decisions e compliance con GDPR, AI Act e legge italiana 132/2025

La possibilità di analizzare 'in chiaro' la logica di discriminazione 'appresa' dall'algoritmo della LLM consente agli esperti, per questo progetto i diabetologi di AMD, di controllare i criteri di funzionamento alla base del modello di machine learning prodotto. Infatti, l'approccio proposto rispetta il paradigma su cui si basa la 'augmented AI', che potenzia e amplia con 'informed decisions' la capacità decisionale del diabetologo, garantendo piena compliance con il GDPR, l'AI Act europeo (Regolamento UE 2024/1689) e la Legge 23 settembre 2025, n. 132 (ricepimento e integrazione nazionale dell'AI Act).

Si sottolinea infatti che, per l'AI act, fra i 'sistemi ad alto rischio' sono inclusi i sistemi di supporto decisionale clinico, algoritmi di diagnostica assistita, sistemi di monitoraggio dei pazienti, software che influiscono su trattamenti terapeutici o chirurgici, i quali necessitano di supervisione umana costante, tracciabilità delle decisioni e valutazioni d'impatto specifiche [37]. Per quanto riguarda la legge 132/25 "Viene anche ribadito, ai commi 2 e 3 della stessa disposizione, la necessità di

sicurezza, trasparenza, correttezza, qualità e appropriatezza dei processi di sviluppo di sistemi di AI per finalità generali, sottolineando la necessità del «rispetto dell'autonomia e del potere decisionale dell'uomo» e dell'assicurazione di possibilità della «sorveglianza e dell'intervento umano» [38].

La stima del valore di eGFR dopo 3 anni, nelle 2 situazioni di utilizzo/non utilizzo di SGLT2

Criteri di selezione e numerosità del campione

Per lo sviluppo dei modelli di regressione finalizzati alla previsione dell'eGFR a 3 anni è stata adottata una metodologia rigorosa per la selezione delle visite e il calcolo delle variabili correlate al fenomeno stimato, distinguendo tra componenti statiche (puntuali) e dinamiche (longitudinali). Lo studio ha utilizzato una coorte retrospettiva di pazienti adulti con diabete di tipo 2 seguiti in ambito specialistico nel periodo 2010-2023 raccolti nel DB Annali, con l'obiettivo di sviluppare due modelli predittivi di regressione lineare, in grado di stimare il valore di eGFR dopo 3 anni a partire da una visita indice (T0): uno in caso di prescrizione di SGLT2 e uno in caso di non prescrizione di SGLT2.

Sono state così composte le due coorti per i due modelli. La coorte SGLT2-SI include i pazienti nei quali, alla visita T0, viene prescritta una terapia con SGLT2 che prosegue per circa 3 anni, senza sospensioni e senza aggiunta di GLP-1 agonista recettoriale. La coorte SGLT2-NO comprende invece i pazienti che non assumono SGLT2 nel periodo osservato. In entrambi i casi, le visite vengono incluse se il paziente presenta almeno 6 anni di follow-up documentato, così da poter stimare l'eGFR 3 anni dopo la visita T0 e calcolare le variabili dinamiche nel periodo T0-3 anni.

Per la coorte SGLT2-SI la visita T0 coincide con la prima prescrizione stabile di SGLT2. Per ciascun paziente si ricostruisce una finestra temporale di circa 3 anni antecedenti T0 per il calcolo delle variabili dinamiche e si identifica una visita di follow-up circa 3 anni dopo T0 (± 150 giorni); il valore di eGFR rilevato a tale visita (eGFR-OUTCOME) costituisce l'outcome del modello di regressione lineare "SGLT2-SI" (**Figura 1**).

Per la coorte SGLT2-NO si considerano i pazienti a cui non è mai stato prescritto SGLT2 e si utilizza una finestra mobile procedendo a ritroso nel tempo (**Figura 2**):

1. Si parte dall'ultima visita (UV) disponibile nel database, da cui si estrae eGFR-OUTCOME, ovvero il valore di eGFR da utilizzare come outcome nel modello di regressione lineare "SGLT2-NO" per quello specifico esempio.
2. Ci si posiziona quindi alla visita collocata circa 3 anni prima di UV, UV-3, da cui si estraggono le variabili statiche di interesse e, per alcune di esse, vengono considerate anche le corrispondenti variabili dinamiche (ad esempio media, deviazione standard, trend), con cui si arricchisce il dataset in input al modello.
3. La finestra mobile viene poi spostata alla visita immediatamente precedente l'ultima visita e si ripetono i passi 1 e 2, generando un nuovo esempio per il modello di regressione "SGLT2-NO".
4. Il processo continua spostando progressivamente la finestra all'indietro nel tempo, finché gli anni che precedono la visita da cui si determina l'outcome (passo 1) risultano pari a 6: a quel punto la procedura si arresta poiché non è più disponibile il range temporale necessario a rilevare eGFR-OUTCOME e a calcolare le variabili dinamiche.

Numerosità del campione

I pazienti che compongono le coorti che alimentano i due modelli sono:

	ESEMPI ¹	PAZIENTI ²	ESEMPI/PAZIENTE
SGLT2-SI	26.743	26.743	1
SGLT2-NO	2.392.901	501.771	4.77

1: numero di visite utilizzate dal modello per il training

2: numero di pazienti che compongono il gruppo definito

Definizione e calcolo delle variabili

L'analisi prevede due tipologie di features per alimentare l'algoritmo:

Variabili Statiche (Snapshot)

Rappresentano la fotografia clinica del paziente in uno specifico momento.

Modalità di rilevazione: vengono estratte puntualmente alla data della singola visita considerata (T0).

Variabili Dinamiche (Trend Longitudinale)

Catturano l'evoluzione nel tempo di alcuni parametri clinici (es. media/stDev/trend).

Finestra di osservazione (Look-back period): per ogni visita target considerata, le variabili dinamiche sono calcolate analizzando TUTTE le visite effettuate nei 3 anni precedenti. Ovvero, per calcolare i trend alla visita T0 si analizzano i dati storici nel periodo che precede T0, ponendo come condizione che esista almeno una visita prima di T0 (altrimenti le variabili dinamiche vengono poste a NULL).

Il processo di selezione utilizzato garantisce che ogni predizione sia basata su dati solidi: le variabili statiche forniscono lo stato attuale, mentre le variabili dinamiche arricchiscono il modello quando esiste una storia clinica sufficiente (almeno 3 anni di storico precedente, con almeno 2 punti dati).

La logica di selezione delle coorti per i due modelli è schematizzata in **Figura 1** e **Figura 2**.

Figura.1 Schema di selezione dei pazienti e dei parametri per le visite che alimentano il modello di stima dell'eGFR nella casistica 'SGLT2-SI'

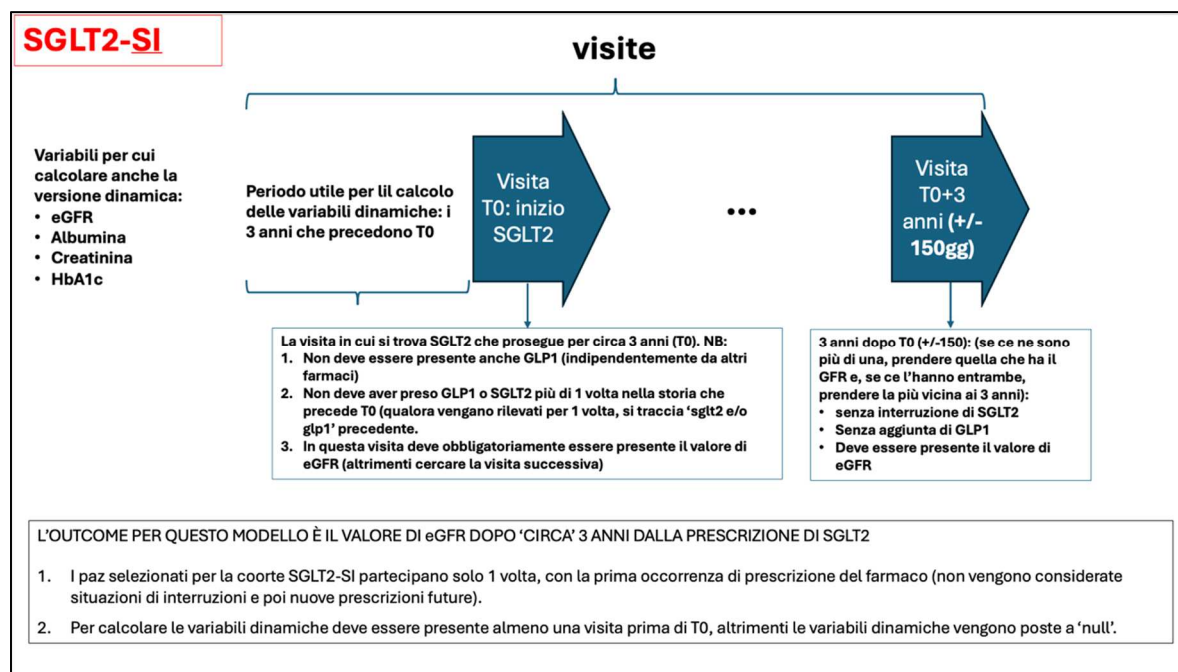
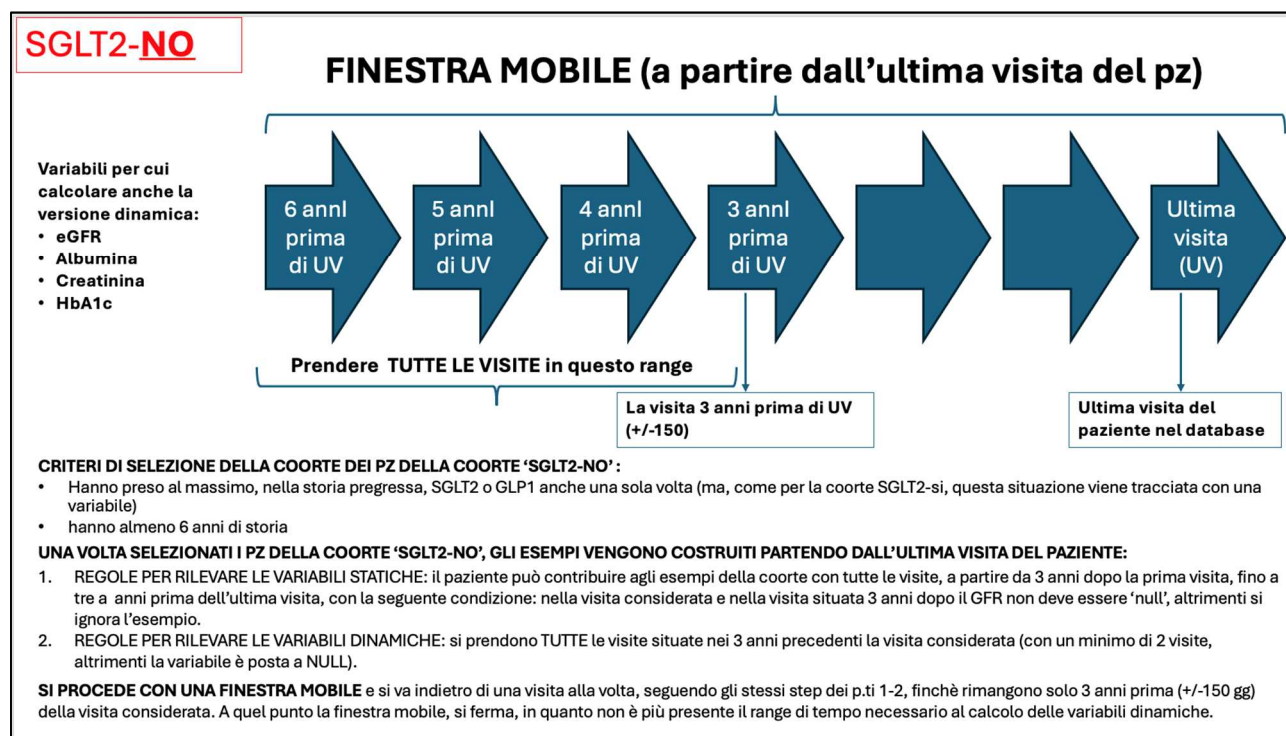


Figura.2 Schema di selezione dei pazienti e dei parametri per le visite che alimentano il modello di stima dell'eGFR nella casistica 'SGLT2-NO'



Modelli di regressione stime eGFR “SGLT2-SI” e “SGLT2-NO”: risultati

Per entrambi i modelli si è proceduto con ottimizzazioni successive:

- Razionalizzando il numero di variabili: i modelli con un numero ridotto di variabili hanno dimostrato le performance migliori.
- Nelle diverse iterazioni, allo scopo di migliorare la precisione e ridurre l'errore, sono stati sviluppati, per entrambe le casistiche (SGLT2-SI/NO), più modelli in base a diversi range di eGFR di partenza (≥ 100 ; 80-99; 60-79; 40-59; < 40).
- Il Root Mean Square Error (RMSE, che misura quanto, in media, le previsioni del modello si discostano dai valori reali, espressa nelle stesse unità della variabile dipendente, nel nostro caso eGFR) si posiziona in un intervallo 5-10.

Per entrambi i modelli sono risultate maggiormente rilevanti le seguenti variabili, in ordine d'importanza:

- eGFR alla visita T0 → variabile nettamente dominante
- di pari rilevanza: età alla visita T0 e albuminuria
- trigliceridi

Questi dati non sono da ritenersi definitivi in quanto le tempistiche stringenti, la complessità e la rilevanza di queste sottoanalisi di stima dell'eGFR rendono opportune ulteriori attività di approfondimento e ottimizzazione.

L'applicazione AMD KDT (Kidney Digital Twin)

NB: la descrizione riportata di seguito rappresenta una prima versione sintetica di specifiche funzionali che verranno ampliate e validate dalla società scientifica.

Scopo dell'applicazione AMD KDT (Kidney Digital Twin)

AMD KDT (Kidney Digital Twin) è una WebApp progettata per migliorare l'appropriatezza terapeutica e gli outcome clinici nelle persone con diabete tipo 2, con un focus specifico sulla preservazione della funzione renale.

L'obiettivo principale è offrire ai diabetologi, ai nefrologi e ai medici coinvolti nella gestione del diabete uno strumento di formazione ed educazione scientifica basato su dati reali e modelli predittivi avanzati.

L'applicazione integra dati clinici storici con algoritmi di machine learning dotati di funzionalità di Explainable AI (XAI), consentendo al medico di esplorare e comprendere i fattori che influenzano la progressione della malattia renale. L'architettura aperta della piattaforma permette inoltre un'evoluzione continua, con l'integrazione futura di ulteriori moduli dedicati alla protezione della funzione renale nel paziente con diabete tipo 2.

Funzioni principali

AMD KDT supporta il clinico in due ambiti chiave:

- Valutazione dell'andamento della funzione renale: la WebApp fornisce una stima del rischio di progressione della malattia renale cronica (MRC), accompagnata da spiegazioni chiare e trasparenti sui principali fattori clinici che hanno contribuito alla previsione.

- Simulazioni what-if in tempo reale: il medico può esplorare scenari alternativi e stimare il valore dell'eGFR a 3 anni in base a diverse strategie terapeutiche, come l'impiego o meno di SGLT2-inibitori.

Utenti e contesto d'uso

- Utenti principali: diabetologi, nefrologi, medici di medicina generale ed altri specialisti con interesse specifico nella gestione della funzione renale nel diabete tipo 2.
- Contesto d'uso: l'applicazione può essere utilizzata in ambulatorio, in reparto o in contesti di formazione medico-scientifica. Non sostituisce il giudizio clinico, ma offre un supporto decisionale basato su evidenze quantitative e facilmente interpretabili.

Conoscenza su cui si basa l'applicazione

L'applicazione AMD KDT integra i modelli sviluppati nell'ambito dell'omonimo progetto, finalizzato a migliorare la gestione clinica delle persone con diabete di tipo 2 e a preservare la funzione renale. I modelli implementati nell'applicazione includono:

- Un modello di classificazione basato sul Machine Learning Trasparente per calcolare il rischio d'insorgenza della MRC entro 3 anni, utilizzando un approccio che garantisce trasparenza e interpretabilità delle decisioni previste dal sistema;
- Un modello di regressione che stima il valore dell'eGFR a 3 anni dalla visita in caso di utilizzo di SGLT2, per aiutare i medici a valutare l'impatto del trattamento;
- Un modello di regressione alternativo che stima il valore dell'eGFR a 3 anni dalla visita in caso di NON utilizzo SGLT2, per confrontare l'efficacia del trattamento rispetto a un gruppo senza terapia farmacologica mirata.

Questi modelli, costruiti su dati reali, permettono di fare previsioni informate e di personalizzare il trattamento per ogni paziente.

Autorizzazioni, amministrazione e requisiti di privacy

Autenticazione

Sono previsti 4 profili di utenti a cui corrispondono 3 gruppi di appartenenza.

- Diabetologo utilizzatore delle funzioni di AMD KDT: ogni diabetologo sarà associato ad un unico centro diabetologico e potrà gestire i dati dei pazienti, invocare i processi predittivi e generare report/estrazioni. Avrà visibilità su tutti i pazienti afferenti al centro di appartenenza.
- Eventuale operatore sanitario che aggiorna i dati dei pazienti (recuperandoli dalle fonti definite da AMD): non ha accesso a report né può eseguire modelli predittivi.
- L'amministratore in grado di creare nuovi centri, creare utenti e assegnarli ad un centro, e abilitarli ad accedere alle funzioni specifiche.
- Un superuser che avrà accesso completo per eseguire modifiche generali sul sistema, con privilegi superiori rispetto agli altri utenti.

Per entrambi i profili, Diabetologo e Operatore sanitario, l'accesso alla WebApp è consentito dopo la fase di autenticazione che avviene con login e password (o con invio di una passkey).

È prevista la gestione di password dimenticate.

L'anagrafica sarà gestita con login e password criptati in una tabella del database; eventuali dati descrittivi aggiuntivi per il medico o l'operatore saranno gestiti con un'altra tabella.

Amministrazione utenti

Da un apposito menu abilitato ai soli utenti del gruppo amministrazione sarà possibile creare, modificare, cancellare utenti (ma non lo storico delle analisi).

Il superuser potrà fare modifiche di tipo 'avanzato'.

Ogni intervento sarà registrato nel registro degli eventi, assieme alla data dell'evento.

Requisiti di sicurezza e privacy

- Gestione degli accessi tramite credenziali nominali.
- Visualizzazione solo dei pazienti per cui l'utente ha autorizzazione, secondo le politiche dell'ente (ruoli, unità operative).

Trattamento dei dati personali in conformità al GDPR e alla normativa nazionale con pseudo-anonimizzazione o anonimizzazione, in base alla decisione della società scientifica. In caso di pseudo-anonimizzazione verrà definito un ente terzo che conserverà i dati di collegamento.

Logging, audit e tracciabilità

L'applicazione dovrà registrare:

- Gli accessi degli utenti;
- Le esecuzioni dei modelli (data/ora, utente, modello, versione del modello);
- Eventuali errori di import o calcolo.

Queste informazioni devono essere esportabili per audit o per documentazione a fini di ricerca.

Funzionalità principali lato utente

Import e preparazione dei dati clinici

Per ciascun paziente l'applicazione deve consentire di:

- Imputare (o importare automaticamente in modalità da definire) i parametri necessari al funzionamento di AMD KDT, tra cui dati anagrafici, storia clinica e principali parametri metabolici, lipidici, pressori, renali e di complicanze del diabete di tipo 2.
- Calcolare, per ciascun indicatore rilevante, gli eventuali parametri dinamici richiesti dai modelli.

La WebApp app dovrà segnalare in modo chiaro se i dati disponibili non sono sufficienti per applicare correttamente uno o più modelli.

Gestione del paziente

L'applicazione deve permettere di:

- ricercare e selezionare un paziente tramite identificativo univoco anonimizzato;
- visualizzare un riepilogo sintetico dei principali dati anagrafici e clinici correlati con lo scopo applicativo.

La creazione e modifica dei dati anagrafici avviene nel sistema clinico sorgente; AMD KDT lavora in sola lettura sui dati importati.

Funzione di calcolo del rischio di MRC entro 3 anni

L'applicazione si basa su un modello di machine learning interpretabile (LLM / XAI), addestrato su dati di coorte reali (DB Annali AMD 2023).

Per ogni paziente selezionato, AMD KDT deve:

- eseguire i modelli con i dati pre-elaborati;
- calcolare la probabilità stimata di insorgenza di MRC entro l'orizzonte temporale definito (3 anni), suddividendo i risultati per outcome MRC-SI / MRC-NO;

- in caso di incertezze nei dati il sistema lo evidenzia (es. risultato da interpretare con cautela per dati incompleti);
- fornire una breve spiegazione testuale standardizzata (“template”) del motivo per cui il paziente è stato classificato a un certo livello di rischio, basata sulle principali regole attivate dal modello.
- fornire i riferimenti bibliografici dalla letteratura scientifica a supporto della previsione effettuata.

Funzione di stima del filtrato glomerulare dopo 3 anni dalla visita

L'applicazione si basa su un modello di regressione addestrato su dati di coorte reali (DB Annali AMD 2023).

Per ogni paziente selezionato, AMD KDT deve:

- eseguire i modelli utilizzando i dati pre-elaborati;
- stimare il valore dell'eGFR dopo l'orizzonte temporale definito (3 anni);
- evidenziare eventuali incertezze nei dati (es. risultato da interpretare con cautela per dati incompleti);
- fornire i riferimenti bibliografici dalla letteratura scientifica a supporto della previsione effettuata.

Bibliografia

- [1] Stevens PE, Ahmed SB, Carrero JJ, Foster B, Francis A, Hall RK, et al. KDIGO 2024 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. *Kidney International* 2024;105:S117–314. <https://doi.org/10.1016/j.kint.2023.10.018>.
- [2] Kalantar-Zadeh K, Jafar TH, Nitsch D, Neuen BL, Perkovic V. Chronic kidney disease. *The Lancet* 2021;398:786–802. [https://doi.org/10.1016/S0140-6736\(21\)00519-5](https://doi.org/10.1016/S0140-6736(21)00519-5).
- [3] Lewis EJ, Hunsicker LG, Clarke WR, Berl T, Pohl MA, Lewis JB, et al. Renoprotective Effect of the Angiotensin-Receptor Antagonist Irbesartan in Patients with Nephropathy Due to Type 2 Diabetes. *N Engl J Med* 2001;345:851–60. <https://doi.org/10.1056/NEJMoa011303>.
- [4] Brenner BM, Cooper ME, De Zeeuw D, Keane WF, Mitch WE, Parving H-H, et al. Effects of Losartan on Renal and Cardiovascular Outcomes in Patients with Type 2 Diabetes and Nephropathy. *N Engl J Med* 2001;345:861–9. <https://doi.org/10.1056/NEJMoa011161>.
- [5] American Diabetes Association Professional Practice Committee, ElSayed NA, McCoy RG, Aleppo G, Balapattabi K, Beverly EA, et al. 11. Chronic Kidney Disease and Risk Management: Standards of Care in Diabetes—2025. *Diabetes Care* 2025;48:S239–51. <https://doi.org/10.2337/dc25-S011>.
- [6] Pozo Garcia L, Thomas SS, Rajesh H, Navaneethan SD. Progress in the management of patients with diabetes and chronic kidney disease. *Current Opinion in Nephrology & Hypertension* 2022;31:456–63. <https://doi.org/10.1097/MNH.0000000000000811>.
- [7] Perkovic V, Jardine MJ, Neal B, Bompont S, Heerspink HJL, Charytan DM, et al. Canagliflozin and Renal Outcomes in Type 2 Diabetes and Nephropathy. *N Engl J Med* 2019;380:2295–306. <https://doi.org/10.1056/NEJMoa1811744>.
- [8] Heerspink HJL, Stefánsson BV, Correa-Rotter R, Chertow GM, Greene T, Hou F-F, et al. Dapagliflozin in Patients with Chronic Kidney Disease. *N Engl J Med* 2020;383:1436–46. <https://doi.org/10.1056/NEJMoa2024816>.
- [9] The EMPA-KIDNEY Collaborative Group. Empagliflozin in Patients with Chronic Kidney Disease. *N Engl J Med* 2023;388:117–27. <https://doi.org/10.1056/NEJMoa2204233>.
- [10] The EMPA-KIDNEY Collaborative Group. Long-Term Effects of Empagliflozin in Patients with Chronic Kidney Disease. *N Engl J Med* 2025;392:777–87. <https://doi.org/10.1056/NEJMoa2409183>.
- [11] Rossing P, Bakris G, Perkovic V, Pratley R, Tuttle KR, Mahaffey KW, et al. Effects of Semaglutide With or Without Concomitant Mineralocorticoid Receptor Antagonist Use in Participants With Type 2 Diabetes and Chronic Kidney Disease: A FLOW Trial Prespecified Secondary Analysis. *Diabetes Care* 2025;48:1878–87. <https://doi.org/10.2337/dc25-0472>.
- [12] Marso SP, Daniels GH, Brown-Frandsen K, Kristensen P, Mann JFE, Nauck MA, et al. Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. *N Engl J Med* 2016;375:311–22. <https://doi.org/10.1056/NEJMoa1603827>.
- [13] Marso SP, Bain SC, Consoli A, Eliaschewitz FG, Jódar E, Leiter LA, et al. Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. *N Engl J Med* 2016;375:1834–44. <https://doi.org/10.1056/NEJMoa1607141>.
- [14] Natale P, Green SC, Tunnicliffe DJ, Pellegrino G, Toyama T, Strippoli GF. Glucagon-like peptide 1 (GLP-1) receptor agonists for people with chronic kidney disease and diabetes. *Cochrane Database of Systematic Reviews* 2025;2025. <https://doi.org/10.1002/14651858.CD015849.pub2>.

- [15]Solis-Herrera C, Triplitt C. Non-steroidal mineralocorticoid receptor antagonists in patients with chronic kidney disease and type 2 diabetes. *Diabetes Obesity Metabolism* 2024;26:417–30. <https://doi.org/10.1111/dom.15327>.
- [16]Bakris GL, Agarwal R, Anker SD, Pitt B, Ruilope LM, Rossing P, et al. Effect of Finerenone on Chronic Kidney Disease Outcomes in Type 2 Diabetes. *N Engl J Med* 2020;383:2219–29. <https://doi.org/10.1056/NEJMoa2025845>.
- [17]Pitt B, Filippatos G, Agarwal R, Anker SD, Bakris GL, Rossing P, et al. Cardiovascular Events with Finerenone in Kidney Disease and Type 2 Diabetes. *N Engl J Med* 2021;385:2252–63. <https://doi.org/10.1056/NEJMoa2110956>.
- [18]Penno G, Solini A, Bonora E, Orsi E, Fondelli C, Zerbini G, et al. Defining the contribution of chronic kidney disease to all-cause mortality in patients with type 2 diabetes: the Renal Insufficiency And Cardiovascular Events (RIACE) Italian Multicenter Study. *Acta Diabetol* 2018;55:603–12. <https://doi.org/10.1007/s00592-018-1133-z>.
- [19]Pugliese G, Solini A, Bonora E, Fondelli C, Orsi E, Nicolucci A, et al. Chronic kidney disease in type 2 diabetes: Lessons from the Renal Insufficiency And Cardiovascular Events (RIACE) Italian Multicentre Study. *Nutrition, Metabolism and Cardiovascular Diseases* 2014;24:815–22. <https://doi.org/10.1016/j.numecd.2014.02.013>.
- [20]Abbafati C, Nieddu L, Monasta L. Measures of type 2 diabetes burden in Italy assessed using the AMD dataset over a twelve year span across the Great Recession. *Sci Rep* 2024;14:4901. <https://doi.org/10.1038/s41598-024-54989-8>.
- [21]Viazzì F, Russo GT, Ceriello A, Fioretto P, Giorda C, De Cosmo S, et al. Natural history and risk factors for diabetic kidney disease in patients with T2D: lessons from the AMD-annals. *J Nephrol* 2019;32:517–25. <https://doi.org/10.1007/s40620-018-00561-3>.
- [22]De Cosmo S, Viazzì F, Pacilli A, Giorda C, Ceriello A, Gentile S, et al. Predictors of chronic kidney disease in type 2 diabetes: A longitudinal study from the AMD Annals initiative. *Medicine* 2016;95:e4007. <https://doi.org/10.1097/MD.0000000000004007>.
- [23]Russo GT, Giandalia A, Lucisano G, Rossi MC, Piscitelli P, Pontremoli R, et al. Prevalence and clinical determinants of rapid eGFR decline among patients with newly diagnosed type 2 diabetes. *European Journal of Internal Medicine* 2024;130:123–9. <https://doi.org/10.1016/j.ejim.2024.07.034>.
- [24]De Boer IH, Khunti K, Sadusky T, Tuttle KR, Neumiller JJ, Rhee CM, et al. Diabetes Management in Chronic Kidney Disease: A Consensus Report by the American Diabetes Association (ADA) and Kidney Disease: Improving Global Outcomes (KDIGO). *Diabetes Care* 2022;45:3075–90. <https://doi.org/10.2337/dci22-0027>.
- [25]Navaneethan SD, Jolly SE, Sharp J, Jain A, Schold JD, Jr MJS, et al. Electronic health records: a new tool to combat chronic kidney disease? *CN* 2013;79:175–83. <https://doi.org/10.5414/CN107757>.
- [26]Jhamb M, Weltman MR, Devaraj SM, Lavenburg L-MU, Han Z, Alghwiri AA, et al. Electronic Health Record Population Health Management for Chronic Kidney Disease Care: A Cluster Randomized Clinical Trial. *JAMA Intern Med* 2024;184:737. <https://doi.org/10.1001/jamainternmed.2024.0708>.
- [27]Jones JL, Lumsden NG, Simons K, Ta'eed A, De Courten MP, Wijeratne T, et al. Using electronic medical record data to assess chronic kidney disease, type 2 diabetes and cardiovascular disease testing, recognition and management as documented in Australian general practice: a cross-sectional analysis. *Fam Med Com Health* 2022;10:e001006. <https://doi.org/10.1136/fmch-2021-001006>.
- [28]Chu Y, Li S, Tang J, Wu H. The potential of the Medical Digital Twin in diabetes management: a review. *Front Med* 2023;10:1178912. <https://doi.org/10.3389/fmed.2023.1178912>.

- [29]Zilich, R. Digital twins in diabetes care: transforming technology into clinical practice. JAMD 2025;28:53. <https://doi.org/10.36171/jamd25.28.1-2.7>.
- [30]De Domenico M, Allegri L, Caldarelli G, d'Andrea V, Di Camillo B, Rocha LM, et al. Challenges and opportunities for digital twins in precision medicine from a complex systems perspective. Npj Digit Med 2025;8:37. <https://doi.org/10.1038/s41746-024-01402-3>.
- [31]Su C, Wang P, Foo N, Ho D. Optimizing metabolic health with digital twins. Npj Aging 2025;11:20. <https://doi.org/10.1038/s41514-025-00211-6>.
- [32]Fatehi M. How Digital Twins Are Revolutionizing Diabetes Care. Digital Health Perspectives. 2025.
- [33]Muselli M, Ferrari E. Coupling Logical Analysis of Data and Shadow Clustering for Partially Defined Positive Boolean Function Reconstruction. IEEE Trans Knowl Data Eng 2011;23:37–50. <https://doi.org/10.1109/TKDE.2009.206>.
- [34]Muselli M. Switching Neural Networks: A New Connectionist Model for Classification. In: Apolloni B, Marinaro M, Nicosia G, Tagliaferri R, editors. Neural Nets, vol. 3931, Berlin, Heidelberg: Springer Berlin Heidelberg; 2006, p. 23–30. https://doi.org/10.1007/11731177_4.
- [35]Verda D, Parodi S, Ferrari E, Muselli M. Analyzing gene expression data for pediatric and adult cancer diagnosis using logic learning machine and standard supervised methods. BMC Bioinformatics 2019;20:390. <https://doi.org/10.1186/s12859-019-2953-8>.
- [36]Muselli M. Extracting knowledge from biomedical data through Logic Learning Machines and Rulex. EMBnet j 2012;18:56. <https://doi.org/10.14806/ej.18.B.549>.
- [37]Cassano G, Tripodi EM. Il regolamento europeo sull'Intelligenza Artificiale. vol. 978-8891670885. Maggioli Editore; 2024.
- [38]Iaselli M. Commentario alla legge sull'intelligenza artificiale. Principi, governance e applicazioni settoriali. Legge 23 settembre 2025, n. 132. EPC editore; 2025.

Realizzato grazie al contributo non condizionante di

