

Decidere per i pazienti incapaci ai tempi dell'IA: etica dell'algoritmo predittivo delle preferenze

Deciding for Incapacitated Patients in the Age of AI:
The Ethics of Preference-Predicting Algorithms

DAVIDE BATTISTI

Assegnista di ricerca, Dipartimento di Giurisprudenza, Università degli Studi di Bergamo; Professore a contratto, Dipartimento di Scienze sociali e politiche, Università degli Studi di Milano.

E-mail: Davide.battisti@unibg.it

ABSTRACT

Gli sviluppi dell'intelligenza artificiale potrebbero consentire la realizzazione di algoritmi per la previsione delle preferenze di pazienti in condizioni di incapacità decisionale. Secondo alcuni, questi strumenti rappresentano un'opportunità per promuovere l'autonomia del paziente. Tuttavia, accettare la loro implementazione comporta un profondo ripensamento delle decisioni sostitutive. Questo articolo analizza criticamente alcune delle più rilevanti questioni bioetiche sottese a tale dibattito.

Artificial intelligence developments could enable the creation of algorithms to predict the preferences of patients in conditions of decisional incapacity. According to some, these tools represent an opportunity to promote patient autonomy. However, their implementation requires a profound reconsideration of surrogate decision-making. This article offers new insight into some of the most relevant bioethical issues underlying this debate.

KEYWORDS

giudizio sostitutivo, intelligenza artificiale, algoritmo predittivo delle preferenze, disposizioni anticipate di trattamento, autonomia

substituted judgment, artificial intelligence, algorithmic preference predictor, advance directives, autonomy

Decidere per i pazienti incapaci ai tempi dell'IA: etica dell'algoritmo predittivo delle preferenze

DAVIDE BATTISTI

1. Introduzione – 2. Le disposizioni anticipate di trattamento e il giudizio sostitutivo – 3. Gli algoritmi predittivi delle preferenze: differenze tecniche ed etiche – 4. Implicazioni etiche se i risultati dell'algoritmo sono affidabili – 4.1. Affidabilità dell'algoritmo predittivo: problemi tecnici ed epistemici – 4.2. Implicazioni etiche se i risultati dell'algoritmo sono affidabili – 4.2.1. Algoritmo predittivo delle preferenze personalizzato: una (nuova) questione di autonomia? – 5. Conclusioni: la necessità del ruolo attivo del paziente e l'IA come supporto a una efficace redazione delle DAT.

1. Introduzione

Gli avanzamenti tecnologici nell'ambito degli algoritmi computazionali, del *machine learning* e dell'intelligenza artificiale (IA) offrono significative opportunità di applicazione in ambito medico. La capacità di questi strumenti di connettere e assemblare informazioni in modo rapido ed efficace può, ad esempio, perfezionare il processo diagnostico, nonché fornire una migliore individuazione dei trattamenti per il paziente nella prospettiva di una medicina sempre più personalizzata. Inoltre, questi strumenti potrebbero essere impiegati per il monitoraggio di sintomi o di informazioni clinicamente rilevanti al fine di gestire patologie complesse, come quelle croniche.

Alla luce di questi rapidi e significativi sviluppi, anche il recente dibattito bioetico si è occupato in modo estensivo delle implicazioni etiche dell'applicazione dell'IA nell'ambito della biomedicina¹. In questa direzione sono state affrontate diverse questioni, tra cui: il rischio che l'IA possa incorporare dei *bias*, reiterando ed esacerbando discriminazioni e iniquità²; i problemi che emergono dall'utilizzo dei cosiddetti "dati storici"³ – che a differenza dei dati sintetici appartengono a persone reali – per allenare gli algoritmi e la discussione relativa ai modelli di consenso informato per la loro raccolta e conservazione⁴; i profili di responsabilità del professionista sanitario in caso di errore della macchina⁵; l'opacità dell'IA e la necessità di uno strumento che sia anche *spiegabile* non solo per assicurare una supervisione adeguata della macchina ed evitare possibili errori, ma anche per scongiurare l'emergenza di un nuovo paternalismo medico⁶; l'impatto dell'IA sulla relazione medico-paziente o paziente-professionista sanitario sia nell'ambito diagnostico e prognostico sia in quello comunicativo⁷ e assistenziale⁸.

Oltre a tali questioni, il dibattito si è concentrato anche sulle questioni etiche che emergono da alcune possibili applicazioni dell'IA in contesti di fine vita nell'ambito del giudizio sostitutivo per i pazienti che hanno perso la capacità decisionale. Infatti, l'IA può essere uno strumento molto performante non solo nella diagnosi di patologie o nella definizione di percorsi terapeutici, ma anche nella predizione delle preferenze dei pazienti. Ciò diventa ancor più rilevante se i

¹ KARIMIAN et al. 2022; MORLEY et al. 2020; ANYANWU et al. 2024; MANNELLI 2022; CNB 2020.

² MURPHY ET AL. 2021; GIOVANOLA, TIRIBELLI 2023; PALAZZANI 2020.

³ FLORIDI 2023.

⁴ SAVULESCU et al. 2024.

⁵ NAIK et al. 2022.

⁶ LORENZINI et al. 2023.

⁷ LEVINE et al. 2024.

⁸ DALTON-BROWN 2020.

pazienti versano in uno stato nel quale non è più possibile per loro decidere in merito a quali trattamenti sottoporsi.

Se da un lato alcuni autori hanno affermato che queste prospettive tecniche possono rappresentare un'inedita opportunità per assicurare e promuovere spazi di autonomia e autodeterminazione per il paziente, dall'altro accettare la possibile implementazione di tali strumenti in contesti tragici e complessi come quelli in esame conduce a un profondo ripensamento del modo in cui vengono attualmente prese le decisioni sostitutive. Ciò rende urgente e necessaria un'attenta analisi critica di tali scenari. In questo contributo si intende offrire una panoramica della letteratura più rilevante sul tema, nonché una discussione delle recenti proposte relative alla possibile implementazione di questi strumenti nella pratica clinica.

Nello specifico, il presente scritto è strutturato come segue: per prima cosa, verranno discusse le modalità con cui vengono attualmente prese le decisioni per i pazienti incapaci, introducendo lo strumento delle disposizioni anticipate di trattamento (DAT); in questo contesto, verrà fatto cenno alla legge italiana che ne regola l'utilizzo, la n. 219/2017. Dopo aver presentato alcuni dati che attestano lo scarso utilizzo delle DAT da parte della cittadinanza, si esporranno le ragioni per le quali potrebbe essere auspicabile l'utilizzo di un algoritmo predittivo delle preferenze. Successivamente, verranno introdotti tre dei più influenti modelli algoritmici che sono stati proposti nel dibattito bioetico – il *Patient Preference Predictor* (PPP), l'*Autonomy Algorithm* e il *Personalized Patient Preference Predictor* (P4) – e ne verranno analizzate le giustificazioni etiche, nonché le critiche che sono state sollevate nei loro confronti. In particolare, si discuteranno due classi di problemi: a) le sfide tecniche ed epistemiche che possono ostacolare il raggiungimento di una predizione affidabile da parte dei modelli algoritmici; b) le implicazioni bioetiche che emergono assumendo che i risultati algoritmici siano effettivamente in linea con le preferenze del paziente. In merito a quest'ultima classe, ci si chiederà quale ruolo e peso normativo dovrebbero avere le previsioni dello strumento in esame. Pertanto, si introdurranno e discuteranno le posizioni della “prevalenza algoritmica moderata” e della “prevalenza algoritmica forte”, interrogandosi sulla possibilità che il giudizio dell'algoritmo debba essere prevalente nell'eventualità di un conflitto con il giudizio dei familiari, del fiduciario o addirittura con quello espresso in precedenza nelle DAT dal paziente stesso. In questa prospettiva, verrà discussa una recente critica a queste posizioni, secondo cui l'algoritmo predittivo non sarebbe in grado di promuovere l'autonomia del paziente a causa di un possibile fraintendimento dei principi etici sottesi allo standard del giudizio sostitutivo. Dopo aver sostenuto che questa critica indebolisce, anche se non in termini definitivi, la posizione della prevalenza algoritmica forte, si concluderà che l'algoritmo predittivo delle preferenze rimane uno strumento problematico sotto vari profili e che per risolvere questi problemi si dovrebbe richiedere un certo coinvolgimento attivo del paziente nell'allenamento dell'algoritmo. Alla luce di ciò, si sosterrà che lo sforzo richiesto al paziente potrebbe essere più adeguatamente impiegato nella redazione e nella raccolta di DAT solide e informate, grazie a uno strumento di IA. In questo contesto, si proporrà l'introduzione delle DAT dinamiche supportate dall'IA.

2. Le disposizioni anticipate di trattamento e il giudizio sostitutivo

In Italia, la possibilità di redigere le DAT è prevista dall'art. 4 della legge n. 219/2017⁹, la quale è il frutto di un percorso avviato nei primi anni 2000 e di un complesso iter parlamentare¹⁰. Grazie

⁹ Per una discussione sugli aspetti bioetici e biogiuridici della legge n. 219/2017, si veda, tra gli altri, BORSELLINO 2018a; D'AMICO M., PELLIZZONE I., LIBERALI B. 2019; DE MARZO 2018; FERRANDO G., BUSATTA L., FRANCESCONI A., BERGONZINI C., MOLASCHI V., VERONESI P., BENCIOLINI P., GIARDINA F., PIZZETTI F.G., CANESTRARI S., MORINO P., VIAFORA C., GIANNINI A., ZAMPERETTI N., PARIS D., GRISTINA G.R., ORSI L. 2018; PALAZZANI 2018.

a questo strumento, una persona può predisporre un documento legalmente vincolante che comunichi al medico e ai familiari i trattamenti a cui intende o non intende sottoporsi in determinate circostanze, qualora si trovasse in uno stato di incapacità. La legge n. 219/2017 prevede, inoltre, la possibile nomina di un fiduciario al quale viene riconosciuta la facoltà di prendere decisioni mediche per il paziente, evitando in questo modo possibili conflitti che potrebbero emergere tra i familiari. Secondo il comma 6 dell'art. 4, affinché le DAT siano vincolanti, queste devono essere redatte per atto pubblico o scrittura privata autenticata e depositate presso un apposito registro del comune di residenza della persona maggiorenne¹¹, o in alternativa presso notai, strutture sanitarie o strutture consolari¹².

Le DAT sono oggi considerate un presidio contro l'accanimento terapeutico e uno strumento standard per il rispetto dell'autonomia del paziente incapace. Sono la risposta a un problema profondo che si era posto già a partire dagli anni '60 del secolo scorso, quando la capacità tecnologica ha reso possibile prolungare la vita in termini biologici anche se a questa non sempre corrisponde un'estensione della vita biografica. Ciò ha fatto emergere in molte persone la preoccupazione di rimanere intrappolate tra le maglie del progresso tecnologico, nel caso in cui venisse meno la possibilità di esprimere le proprie volontà su come e per quanto tempo essere trattate alla fine della vita¹³.

Ciononostante, le DAT rimangono uno strumento ancora poco utilizzato. Evidenza di ciò è fornita dall'ultimo rapporto dell'Osservatorio permanente sulle DAT dell'Associazione Coscioni¹⁴. Nei comuni italiani che hanno risposto all'indagine dell'Osservatorio (6.096), risultano depositate solamente 230.940 DAT¹⁵. Sebbene il dato sia in forte aumento rispetto al 2022 (+52,5%), rimane lampante che le persone che sottoscrivono le DAT sono ancora una esigua minoranza rispetto alla totalità delle persone che potrebbero redigerle e depositarle.

Questo fenomeno può essere dovuto a diversi fattori, legati sia alla specificità della situazione italiana sia a considerazioni più generali. In primo luogo, nel contesto italiano, le DAT sono state formalmente introdotte meno di dieci anni fa e quindi potrebbe essere necessario un orizzonte temporale più lungo per apprezzare un maggiore utilizzo di questo strumento. In merito a questa osservazione, va tuttavia riconosciuto che già dal 2006 il codice di deontologia medica prevedeva il dovere di tenere in considerazione eventuali volontà precedentemente espresse da un paziente non più competente¹⁶; inoltre, un ampio dibattito pubblico su questo tema era già iniziato a partire dal 1999, quando Beppino Englaro richiese al Tribunale di Lecco l'interruzione dei trattamenti vitali per la figlia Eluana, la quale versava in stato vegetativo. In secondo luogo, le modalità di deposito delle DAT potrebbero disincentivare molte persone a intraprendere questa procedura¹⁷. Infine, è opportuno riconoscere che molte persone – soprattutto se non affette da patologie gravi – tendono a non voler parlare della propria morte o di situazioni in cui potrebbero trovarsi in uno stato di incapacità; molti medici sono altresì reticenti a sollevare l'argomento e ad aiutare i pazienti a pensare concretamente a come potrebbe avvenire la loro morte¹⁸. Queste ultime considerazioni assu-

¹⁰ BORSELLINO 2018b; LENZI 2018.

¹¹ L'art. 4 della legge n. 219/2017 non prevede la possibilità di redigere le DAT. In BALDINI 2019 vengono evidenziati i possibili effetti discriminatori, i quali si estendono anche alla categoria degli incapaci. Per una disamina delle implicazioni sui minori di questa legge si veda DI COSTANZO 2019.

¹² Nel 2020 è stata istituita la Banca dati nazionale delle DAT con lo scopo di raccolta dei documenti depositati, la tempestiva gestione in caso di rinnovo, modifica o revoca e la piena accessibilità della DAT su tutto il territorio nazionale. Questo aspetto riesce a mitigare una delle limitazioni sottolineate da CAPRON 2010, ovvero l'indisponibilità, in alcuni contesti, delle DAT anche se queste erano state redatte e depositate.

¹³ CAPRON 2010.

¹⁴ OSSERVATORIO PERMANENTE SULLE DAT 2024.

¹⁵ A questo dato andrebbero aggiunte le DAT depositate nei comuni che non hanno risposto alle richieste e quelle depositate presso notai, strutture sanitarie e uffici consolari a cui l'Osservatorio DAT non ha accesso.

¹⁶ REICHLIN 2021.

¹⁷ BORSELLINO 2019.

¹⁸ CAPRON 2010.

mono particolare rilevanza se si considera che la legge n. 219/2017 non impone al medico l'obbligo di prospettare ai pazienti la possibilità di redigere le DAT¹⁹. In generale, la scarsa attitudine a parlare di morte e di fine vita potrebbe quindi disincentivare il ricorso a questo strumento. Questi aspetti evidenziano un atteggiamento antropologico che trascende i confini nazionali. Lo scarso utilizzo delle DAT, infatti, non riguarda solo l'Italia: ad esempio, nel 2017 è stato osservato che solo una persona su tre negli Stati Uniti aveva redatto una disposizione anticipata²⁰, sebbene la prima iniziativa in uno Stato federale sia avvenuta già nel 1976 in California.

Un ridotto ricorso alle DAT conduce inevitabilmente a dover gestire un elevato numero di casi in cui il paziente che versa in una condizione di incapacità non ha lasciato indicazioni. A questi se ne devono aggiungere anche altri in cui, seppur la DAT sia stata redatta, questa non risulta chiara o non comprende la situazione specifica in cui versa il paziente. In tutti questi casi – che vicino alla fine della vita dei pazienti rappresenterebbero quasi il 70% del totale²¹ – l'onere della decisione viene fatto ricadere su un decisore sostituto, generalmente un familiare, il quale ci si attende che prenda una decisione che rispetti quanto più possibile le preferenze che il paziente avrebbe espresso qualora quest'ultimo avesse potuto esprimersi. In questo senso, si fa spesso riferimento al cosiddetto “standard del giudizio sostitutivo” (*substituted judgment standard*)²². Ci si affida ai familiari perché si suppone che questi riescano a prendere una decisione rispettando due criteri che Hubbard e Greenblum definiscono come il “criterio del vantaggio epistemico” e il “criterio della fedeltà”²³. Per quanto riguarda il primo, in virtù della relazione speciale con il paziente, i familiari sarebbero in una posizione epistemicamente privilegiata per identificare quale preferenza avrebbe il paziente se si trovasse nelle condizioni di poter decidere per sé²⁴. Il secondo criterio richiede che la scelta rispetti le preferenze del paziente e non quelle del decisore sostituto; in questa prospettiva, è verosimile aspettarsi che il familiare sia pronto ad agire unicamente nell'interesse del proprio caro.

Nonostante l'idea che i familiari possano adeguatamente agire come decisori sostitutivi di pazienti incapaci sia attraente e in linea con le nostre intuizioni, le evidenze empiriche disponibili suggeriscono il contrario. Diversi studi, infatti, hanno constatato come le persone chiamate a pronunciarsi per conto del paziente incapace non predicono le preferenze di quest'ultimo in termini superiori alla probabilità di base²⁵. Altri studi hanno persino dimostrato che la probabilità che i familiari identifichino il trattamento desiderato dal paziente in casi complessi – come, ad esempio, quando non è chiaro quale trattamento sia quello clinicamente indicato – è solo di poco superiore al caso²⁶. In altre parole, non è vero che i familiari *qua* decisori sostitutivi soddisfano il criterio del vantaggio epistemico. Inoltre, i familiari spesso esperiscono ansia e stress quando viene chiesto loro di prendere una decisione relativa alla vita o alla morte dei propri cari;

¹⁹ Si ringrazia un revisore anonimo per aver sollevato questo aspetto. Per una discussione critica sulla fonte di informazione per il possesso di adeguate informazioni mediche per redigere le DAT nell'ambito della legge n. 219/2017, si veda BALDINI 2019.

²⁰ YADAV et al. 2017.

²¹ JOHN et al. 2020.

²² Lo standard del giudizio sostitutivo non va confuso con un altro standard che viene spesso utilizzato nell'ambito dei pazienti incapaci, che è quello del miglior interesse (*best interest standard*). Spesso non è possibile determinare le preferenze autonome del paziente: ad esempio, quando non si hanno informazioni sufficienti riguardanti le preferenze del paziente o quando il paziente non è mai stato competente, e. g., pazienti con gravi forme di disabilità cognitive, feti, neonati o bambini che non hanno ancora sviluppato la capacità di prendere delle scelte autonome. In questi scenari, il decisore deve determinare «il massimo beneficio netto probabile tra le opzioni disponibili, attribuendo diversi pesi agli interessi che il paziente ha in ciascuna opzione, bilanciati rispetto ai rischi, agli oneri o ai costi intrinseci» (BEAUCHAMP, CHILDRESS 2019, 56). In questo articolo, si considereranno solo le persone che hanno *perso* la capacità di esprimere le proprie preferenze e quindi si escludono le categorie sopra descritte.

²³ HUBBARD, GREENBLUM 2020.

²⁴ BRUDNEY 2009; BUCHANAN, BROCK 1990.

²⁵ SHALOWITZ et al. 2007; HOUTS et al. 2002; SMUCKER et al. 2000.

²⁶ CIROLDI et al. 2007; SHALOWITZ et al. 2006.

questo potrebbe incidere sulla qualità della decisione. Non è dunque del tutto certo che i familiari siano in grado di isolare adeguatamente le preferenze del paziente per promuoverle attraverso la decisione sostitutiva e, pertanto, di agire in linea con il criterio della fedeltà²⁷.

3. Gli algoritmi predittivi delle preferenze: differenze tecniche ed etiche

Le considerazioni presentate nel paragrafo precedente – insieme agli sviluppi dei sistemi algoritmici computazionali, di *machine learning* e di IA – hanno condotto diversi autori a proporre l'introduzione di un predittore di preferenze algoritmico per i pazienti incapaci²⁸. In questa sede vengono analizzate tre proposte, presentate in ordine cronologico dalla meno recente alla più recente: il *Patient Preference Predictor* (PPP), l'*Autonomy Algorithm* e il *Personalized Patient Preference Predictor* (P4). Sebbene vi siano delle differenze moralmente rilevanti tra le proposte, le quali verranno sottolineate alla fine del paragrafo, si conviene che la loro diversità sia dovuta principalmente a un differente orizzonte di sviluppo tecnologico di riferimento.

Proposto inizialmente da Rid e Wendler²⁹, il PPP è un algoritmo computazionale basato sui risultati di un sondaggio volontario in cui a un campione di persone, rappresentativo dal punto di vista sociodemografico, viene chiesto di esprimere le proprie preferenze in diverse situazioni mediche ipotetiche in cui non sarebbero più competenti. Attraverso questi dati, l'algoritmo predittivo potrebbe inferire le preferenze di un paziente che non ha partecipato al sondaggio e si trova in una situazione di incapacità. Età, livello di istruzione, genere e altri dati sociodemografici possono infatti avere un ruolo predittivo in merito alle preferenze del paziente e quindi determinare la scelta che più è in linea con quanto egli avrebbe scelto qualora fosse stato competente. Questo strumento potrebbe supportare i familiari e i clinici nel processo decisionale condiviso, offrendo almeno due benefici: a) la decisione risulterebbe più coerente con le preferenze del paziente; b) gli attori coinvolti verrebbero, almeno in parte, sollevati dall'onere della decisione.

Tuttavia, anche assumendo che il PPP sia generalmente affidabile, questo sarebbe in grado di offrire previsioni solamente se ci si attende che le preferenze del paziente siano in linea con quelle del suo gruppo sociodemografico. In tutti gli altri casi, il PPP sarebbe uno strumento inefficace. Al fine di ovviare a questo problema, Lamanna e Byrne propongono l'*Autonomy Algorithm*³⁰, una versione dell'algoritmo predittivo che integra i dati sociodemografici con altri per offrire delle previsioni più accurate, anche in casi più complessi. Gli autori propongono di allenare l'algoritmo con le cartelle cliniche di un vasto numero di persone – previa disponibilità e autorizzazione –, la cartella clinica personale del paziente incapace e i dati provenienti dai suoi profili social. In una difesa più sistematica di questo modello, Hubbard e Greenblum sottolineano l'importanza dei dati personali dei profili social, riportando evidenze della loro efficacia predittiva³¹: già nel 2013, Kosinski e colleghi hanno dimostrato come i *like* lasciati su Facebook potessero essere utilizzati per predire in modo accurato dati personali sensibili come l'orientamento sessuale, l'etnia, il credo religioso e quello politico, i tratti della personalità, l'intelligenza, la felicità e l'età³². Nella stessa direzione, un altro studio ha dimostrato che l'accuratezza predittiva di un algoritmo riguardo ad alcuni aspetti della personalità e basato sui *like* di Facebook supera addirittura quella dei familiari e delle persone più vicine³³. Le previsioni algoritmiche della

²⁷ HUBBARD, GREENBLUM 2020.

²⁸ EARP et al. 2024; HUBBARD, GREENBLUM 2020; LAMANNA, BYRNE 2018; FERRARIO et al. 2023.

²⁹ RID, WENDLER 2014.

³⁰ LAMANNA, BYRNE 2018.

³¹ HUBBARD, GREENBLUM 2020.

³² KOSINSKI et al. 2013.

³³ YOUYOU et al. 2015.

personalità e delle preferenze delle persone sulla base dell'utilizzo dei social network stanno diventando sempre più accurate e raffinate³⁴. Questo progresso è alimentato sia dall'avanzamento delle capacità tecnologiche sia dall'aumento della mole di informazioni disponibili: secondo Statista, nel 2024 le persone hanno passato circa 143 minuti al giorno sui social³⁵, i quali oggi offrono esperienze di fruizione che non si riducono a quelle considerate dalla letteratura già menzionata. È plausibile che i risultati algoritmici futuri saranno ulteriormente affinati, integrando ai già menzionati *like* anche informazioni relative al tempo trascorso visionando alcune tipologie rilevanti di video o *reel*, ai profili che le persone seguono e così via.

L'ultima proposta che qui si presenta è quella recentemente delineata da Earp e colleghi ed è denominata *Personal Patient Preference Predictor* (P4)³⁶. Sebbene sia molto simile all'*Autonomy Algorithm*, il P4 è differente nella misura in cui viene attribuita ancor più importanza ai dati personali del paziente rispetto a dataset più generali. Inoltre, il P4 applica strumenti di *large language model* alla predizione delle preferenze dei pazienti incapacitati. Un *large language model* è un sistema di IA progettato per generare testi ed è in grado di essere allenato per svolgere una varietà di compiti. Gli autori suggeriscono di allenare un *large language model* (*fine tuning*) con i dati personali del paziente al fine di creare un cosiddetto gemello digitale, cioè un sistema testuale che emula accuratamente i *pattern* di comportamento, lo stile comunicativo e le modalità decisionali del paziente. I dati personali che gli autori considerano per il processo di *fine tuning* dovranno essere dei testi redatti dal o descrittivi il paziente, anche se ci si attende che grazie al rapido avanzamento tecnologico potranno essere utilizzati anche altri formati, come ad esempio contenuti video o registrazioni vocali. I dati includeranno, ad esempio, cartelle cliniche elettroniche e biobanche, risposte a questionari medici o esperimenti di scelta per individuare i valori abbracciati dall'individuo mentre era competente, ma anche interventi su blog, siti web o altri scritti pubblicati, attività sui social media ed e-mail. Grazie al P4, i familiari e i clinici potrebbero addirittura interrogare in tempo reale il risultante gemello digitale per risalire a quale potrebbe essere la più probabile preferenza per il trattamento in qualsiasi crisi sanitaria.

In generale, da un punto di vista bioetico, l'algoritmo predittivo viene difeso invocando il principio di autonomia nella misura in cui tale principio è sovrapponibile al rispetto delle preferenze – non esplicitamente espresse in una DAT – che il paziente avrebbe avuto in quella data circostanza. Questo strumento sembra infatti soddisfare entrambi i criteri dello standard del giudizio sostitutivo. L'algoritmo predittivo si troverebbe in una condizione di vantaggio epistemico in relazione alle preferenze del paziente. Inoltre, mettendo tra parentesi la controversa possibilità che l'algoritmo predittivo venga programmato per soddisfare le preferenze del professionista sanitario, della struttura o del sistema sanitario, o dei familiari, nelle versioni proposte – e in linea con quanto enunciato dalla legge n. 217/2019 e dall'articolo 32 della Costituzione – questo strumento dovrebbe essere concepito *solo* per soddisfare le preferenze del paziente.

Si noti che l'evoluzione delle proposte degli algoritmi predittivi ha permesso di evitare alcune insidiose critiche di cui soffrivano le prime versioni. Si è già fatto cenno al fatto che il PPP non può essere applicato nei casi in cui ci si attende che il paziente si discosti, in termini di valori e preferenze, dalla sua classe sociodemografica di riferimento. A prescindere da ciò, secondo alcuni autori, il PPP minerebbe comunque il principio di autonomia, anziché promuoverlo. L'idea alla base di questo argomento è che utilizzando il giudizio algoritmico non si starebbero considerando i *valori* e le *ragioni* per cui una persona possiede una certa preferenza, la quale verrebbe *causalmente* determinata dalla mera appartenenza a una classe sociodemografica³⁷. In altri termini, al fine di osservare una scelta autonoma non sarebbe sufficiente individuare la preferenza corretta, ma anche le

³⁴ CHRISTIAN et al. 2021.

³⁵ https://www.statista.com/statistics/433871/daily-social-media-usage-worldwide/?utm_source=chatgpt.com.

³⁶ EARP et al. 2024.

³⁷ SHARADIN 2018; ONEIL 2022.

motivazioni attraverso cui si arriva a tale preferenza. L'algoritmo predittivo in questa forma rischierebbe quindi di ridurre la soggettività del paziente ai suoi dati, aprendo a fenomeni di *datafication* e oggettivazione del paziente. Questa critica si applica al PPP, ma non all'*Autonomy Algorithm* e al P4. Infatti, attraverso la considerazione dei dati personali prodotti dall'individuo è possibile farsi guidare da informazioni – quando queste sono sufficientemente esplicite e non meramente inferite – che sono state causalmente determinate da convinzioni, valori e ragioni del paziente e non da informazioni sociodemografiche o che appartengono ad altre persone. Alla luce di ciò, il rischio di *datafication* e di oggettivazione paventato poc'anzi sembrerebbe piuttosto limitato.

Tuttavia, le proposte più avanzate di algoritmo predittivo delle preferenze soffrono di problemi bioetici che richiedono un'attenta valutazione; nel prossimo paragrafo verranno discussi, considerando principalmente le versioni di predittori algoritmici che si basano in modo primario su dati personali e non meramente su informazioni sociodemografiche.

4. Questioni bioetiche degli algoritmi predittivi basati su dati personali

Per discutere delle questioni bioetiche sollevate dal possibile utilizzo dell'algoritmo predittivo è utile operare una distinzione tra due classi di critiche, entrambe rilevanti: la prima classe riguarda i limiti tecnici ed epistemici che possono ostacolare questo strumento nel raggiungimento di una predizione affidabile; la seconda classe comprende le questioni etiche che emergono ipotizzando che la previsione algoritmica possa effettivamente riflettere le preferenze del paziente.

4.1. Affidabilità dell'algoritmo predittivo: problemi tecnici ed epistemici

L'algoritmo predittivo delle preferenze personalizzato è uno strumento ipotetico che non è ancora stato né programmato né implementato; pertanto, non è ancora del tutto chiaro se possa effettivamente funzionare. Potrebbero infatti emergere dei problemi tecnici che ne minerebbero l'efficacia. In questo senso, alcuni autori sono scettici riguardo al funzionamento dell'algoritmo predittivo, sottolineando che i dati rilevanti per la formulazione di preferenze attendibili sarebbero fondamentalmente scarsi³⁸, soprattutto se si considerano quelli potenzialmente più affidabili, ovvero le risposte dirette a sondaggi sulle decisioni in condizioni di incapacità. La scarsità di informazioni rilevanti potrebbe impedire all'algoritmo predittivo di generare risultati attendibili. Inoltre, l'inattendibilità dei risultati potrebbe essere aggravata dai problemi strutturali di cui soffrono i *large language models*, su cui si basa il P4: quando non riescono a formulare una risposta corretta sulla base dei dati a disposizione, questi sistemi possono generare le cosiddette allucinazioni dell'IA, ovvero *output* plausibili ma non corrispondenti a una verità oggettiva e dimostrabile. Questo assume ancor più rilevanza se si considera che spesso gli strumenti di IA sono opachi e quindi diventerebbe difficile comprendere la relazione tra gli *input* (le informazioni rilevanti) e gli *output* del sistema (la supposta preferenza del paziente). Ulteriori problemi tecnici derivano dal fatto che gli *output* dei *large language models* possono essere influenzati dal modo in cui viene formulata la richiesta (*prompting*): in altre parole, è possibile che a seconda del modo di scrivere la richiesta al gemello digitale da parte dei medici o dei familiari, il dispositivo risponda in modo diverso, alimentando dubbi sulla sua affidabilità, perlomeno nel breve periodo³⁹.

Al di là delle questioni tecniche relative all'affidabilità del sistema, è importante considerare anche l'affidabilità delle preferenze su cui viene allenato il predittore algoritmico. In aggiunta ai problemi di scarsità già menzionati in precedenza, è possibile notare che le preferenze rivelate da-

³⁸ BLUMENTHAL-BARBY et al. 2024; RZEPINŃSKI et al. 2024.

³⁹ SHARADIN 2024.

gli individui potrebbero essere differenti da quelle effettive, soprattutto considerando il contesto in cui vengono espresse. Ad esempio, sui social network, le persone potrebbero voler proiettare un'immagine di sé diversa da quella reale o potrebbero semplicemente esprimere preferenze incoerenti⁴⁰. In termini più generali, è poi dubbio che le persone possano esprimere delle preferenze affidabili e stabili nel tempo, ovvero prendere una decisione “ora per allora”; inoltre, le persone sane potrebbero sottostimare la loro qualità della vita nel caso si trovassero in condizioni di disabilità⁴¹. Allenato sulla base di queste informazioni, l'algoritmo potrebbe rispecchiare previsioni errate, non in linea con quanto avrebbe davvero voluto il paziente in quelle determinate condizioni.

È possibile offrire qualche replica alle osservazioni menzionate sopra. Se da una parte è importante riconoscere le difficoltà tecniche attualmente incontrate da uno strumento di questo tipo, è verosimile che gli avanzamenti tecnologici nel futuro prossimo possano risolvere, per lo meno in parte, i problemi sopradescritti. Per quanto riguarda la scarsità dei dati, si deve tenere conto che la nostra quotidianità è sempre più caratterizzata da una commistione con strumenti digitali, i quali potrebbero non solo raccogliere delle informazioni sul nostro conto da blog, e-mail o post sui social network, ma anche registrare le nostre conversazioni e interazioni quotidiane⁴². In una prospettiva meno ipotetica e controversa, potrebbe essere possibile registrare le interazioni con i professionisti sanitari e considerarle per il *fine tuning* del gemello digitale.

Per quanto riguarda il problema della possibile contraddittorietà e inaffidabilità delle preferenze del paziente, si può replicare che il paziente stesso potrebbe decidere di selezionare quali contenuti considerare per il *fine-tuning* del gemello digitale. Inoltre, si potrebbe decidere di escludere contenuti troppo vecchi, riducendo così il rischio di includere preferenze contraddittorie o espressioni isolate ed estemporanee in aperto contrasto con l'insieme delle altre preferenze registrate.

In ogni caso, si potrebbe osservare che le critiche mosse all'algoritmo predittivo riflettono un doppio standard ingiustificato nella valutazione dell'affidabilità delle preferenze predette dall'IA e quelle riportate dalle DAT. Anche quest'ultimo strumento, infatti, è stato ampiamente criticato per problematiche simili⁴³. Pertanto, sarebbe discutibile sostenere che le DAT sono affidabili e appropriate, mentre si nega la stessa affidabilità all'algoritmo predittivo per ragioni che, almeno in alcune circostanze, valgono anche per le DAT. In termini generali, se si ammette che l'algoritmo possa talvolta fallire nel predire preferenze pienamente allineate a quelle che il paziente avrebbe espresso se avesse potuto farlo – a causa di problemi tecnici o dell'inaffidabilità delle preferenze rivelate –, si deve riconoscere allo stesso tempo che gli strumenti attualmente disponibili, come il giudizio sostitutivo o le DAT, presentano anch'essi problemi nel prevedere le reali preferenze del paziente. Pertanto, ai fini dell'implementazione del predittore algoritmico, lo standard richiesto non dovrebbe essere una perfetta affidabilità nella previsione, ma semplicemente una maggiore affidabilità rispetto agli strumenti attualmente in uso per prendere decisioni per pazienti incapaci.

Tuttavia, si riconosce in questa sede che, per raggiungere tale grado di affidabilità, sarà verosimilmente necessario che il paziente riveda periodicamente e in modo attivo alcune informazioni, decidendo se includere o meno determinate fonti per il *fine tuning* dell'algoritmo. Inoltre, è auspicabile che le informazioni per questo processo siano utilizzate a seguito del consenso informato del paziente⁴⁴. Queste considerazioni e l'imprescindibile ruolo attivo del paziente per un'adeguata applicazione dello strumento saranno rilevanti per la riflessione conclusiva che si intende proporre nel paragrafo 5.

⁴⁰ ANNONI 2024.

⁴¹ CROCKER et al. 2021.

⁴² Non si intende che questa prospettiva sia desiderabile, ma che potrebbe essere tecnicamente possibile.

⁴³ REICHLIN 2021.

⁴⁴ EARP et al. 2024.

4.2. Implicazioni etiche se i risultati dell'algoritmo sono affidabili

Il dibattito bioetico non si esaurisce con la valutazione dell'affidabilità dei risultati prodotti dall'algoritmo predittivo. Al contrario, è proprio l'eventualità che tali risultati siano effettivamente in linea con le preferenze del paziente che ha il potenziale di rivoluzionare il modo in cui vengono prese le scelte per le persone incapaci.

Come detto in precedenza, i familiari vengono coinvolti nella decisione sostitutiva in virtù del fatto che soddisfano il criterio del vantaggio epistemico e il criterio della fedeltà. Tuttavia, si è notato come in realtà i familiari non riescano a identificare adeguatamente la preferenza del paziente, motivo per il quale è stato proposto l'algoritmo predittivo, che sembrerebbe avere un maggior vantaggio epistemico rispetto ai familiari. Ciò solleva un tema bioetico delicato e controverso inerente all'effettivo ruolo del predittore algoritmico nella decisione sostitutiva. Sebbene la maggior parte degli autori concepisca l'algoritmo predittivo come un mero supporto ai familiari e ai clinici per raggiungere una decisione relativa al paziente incapace, ci si potrebbe spingere oltre e chiedersi: perché l'IA dovrebbe essere utilizzata *solo* come strumento di supporto invece di avere l'ultima parola in queste decisioni? Proprio in virtù del maggiore vantaggio epistemico dell'IA, alcuni autori hanno sostenuto che, quando un paziente adulto privo di capacità decisionale (a) non ha nominato un fiduciario o (b) non ha redatto una DAT che specifichi la sua preferenza di trattamento⁴⁵, il giudizio dell'algoritmo predittivo dovrebbe avere la precedenza su quello della famiglia, in caso di conflitto. In questa sede si definisce questa posizione della "prevalenza algoritmica moderata", poiché il giudizio dell'IA prevale su quello del familiare ma non su quello del fiduciario e della DAT.

Alla posizione della prevalenza algoritmica moderata possono essere mosse almeno due obiezioni: in primo luogo, dare la precedenza al giudizio sostitutivo dell'IA minerebbe la relazione speciale tra familiare e paziente, aumentando le probabilità che questo processo sia stressante e ansioso. In secondo luogo, il paziente stesso potrebbe preferire che le sue preferenze venissero rispettate, ma anche che ci fosse un certo coinvolgimento dei familiari. È verosimile che, per alcune persone, sia prioritario il coinvolgimento della famiglia nella decisione, soprattutto considerando che molti non hanno una preferenza definita riguardo ai trattamenti a cui sottoporsi in queste circostanze, come visto sopra. In questo senso, si fa riferimento non solo a delle preferenze di *primo livello*, relative a come un paziente vorrebbe essere trattato, ma anche a delle preferenze di *secondo livello*, relative a come effettivamente la decisione deve essere presa⁴⁶.

Si ritiene che queste due obiezioni non invalidino la posizione della prevalenza moderata. Per quanto riguarda la prima, di fronte a un conflitto tra la preferenza del paziente e gli interessi speciali della famiglia è verosimile che l'autonomia del paziente possa avere una certa priorità, qualora si possa prevedere che il paziente abbia delle preferenze ben definite. In altre parole, l'appello alla relazione speciale risulta troppo debole per affermare una certa priorità lessicale di quest'ultima sulla preferenza del paziente. Se così non fosse, dovremmo ammettere che in alcuni casi l'autonomia del paziente debba essere subordinata alla preferenza dei familiari, anche se il paziente è in grado di intendere e di volere⁴⁷. Questa ipotesi, anche alla luce della normativa vigente e l'impianto costituzionale, risulta davvero poco plausibile.

Per quanto riguarda la seconda obiezione, è opportuno notare che il predittore algoritmico personalizzato, soprattutto nella versione proposta da Earp e colleghi, potrebbe fornire risposte complesse che non si limitano a considerare solo le preferenze di primo livello. Infatti, il P4 potrebbe includere anche le preferenze del paziente riguardo al coinvolgimento desiderato dei fa-

⁴⁵ HUBBARD, GREENBLUM 2020.

⁴⁶ EARP et al. 2024.

⁴⁷ HUBBARD, GREENBLUM 2020.

miliari⁴⁸, così come l'importanza che vi attribuisce. Ad esempio, le preferenze relative ai trattamenti ricevuti potrebbero essere più o meno rilevanti rispetto a quelle sul coinvolgimento familiare. In questo modo, sarebbe possibile evitare di violare le preferenze di secondo livello quando queste possono essere rispettate con il supporto del predittore algoritmico. La prevalenza algoritmica moderata sembra quindi salva dalle due obiezioni appena discusse.

Tuttavia, a questo punto si può rilevare un problema più strutturale per questa posizione. Si può infatti osservare che le argomentazioni suggerite per difendere la prevalenza algoritmica moderata sembrano trovare un'applicazione che va oltre i confini proposti dai suoi sostenitori. In particolare, ci si può chiedere perché l'argomento della prevalenza algoritmica debba essere limitato ai soli i casi in cui il paziente non abbia redatto le DAT o nominato un fiduciario. Se tutto ciò che importa è individuare chi (o cosa) riesca a soddisfare in modo più adeguato il criterio del vantaggio epistemico e quello della fedeltà, si potrebbe osservare che l'algoritmo predittivo può risultare superiore anche al fiduciario e, in alcune circostanze, alle DAT. Come già osservato in relazione ai familiari, il fiduciario potrebbe essere soggetto ad ansie e preoccupazioni o potrebbe aver dimenticato quanto gli è stato detto dal paziente. Inoltre, per quanto riguarda le DAT, sebbene il nostro ordinamento non preveda scadenza per questo documento⁴⁹, una disposizione redatta "ora per allora" incontra delle difficoltà a essere propriamente attendibile: ad esempio, le persone potrebbero nel tempo cambiare i propri valori e idee in merito al fine vita, senza tuttavia modificare il documento. In questo senso, si potrebbe affermare che le DAT sarebbero sì uno strumento importante per informare la decisione – certamente non ridondante, perché comunicherebbero una sensibilità peculiare da parte dell'individuo in merito a temi complessi come le decisioni in contesti di incapacità – ma non sufficiente. L'algoritmo predittivo, invece, forte della grande mole di dati personali del paziente tracciati nel corso del tempo, potrebbe paradossalmente avere un vantaggio epistemico rispetto alle DAT stesse, soprattutto se queste sono state redatte molti anni prima e in termini vaghi. A tal proposito, è opportuno riconoscere che spesso le DAT vengono redatte usando un linguaggio poetico anziché preciso, aprendo più dubbi di quelli che si intendeva dissipare con questo strumento.

Pertanto, non vi sarebbero ragioni per escludere l'uso del predittore algoritmico anche in presenza di un'esplicita volontà del paziente, ipotesi esclusa dai sostenitori della prevalenza algoritmica moderata. Il predittore rappresenterebbe uno strumento imprescindibile per promuovere l'autonomia del paziente e, almeno in alcune circostanze di conflitto tra le DAT e la previsione algoritmica, quest'ultima dovrebbe avere la precedenza. Alla luce di ciò, la posizione della prevalenza moderata sembrerebbe irrimediabilmente *ad hoc* e le ragioni utilizzate dai suoi sostenitori condurrebbero verso l'accettazione di una diversa posizione, che viene qui definita come "prevalenza algoritmica forte".

È possibile avanzare delle critiche alla prevalenza algoritmica forte. In primo luogo, e in termini più circoscritti, si può osservare che l'esplicita nomina del fiduciario non sempre viene fatta per soddisfare una preferenza di primo livello, ma *solamente* per soddisfare una preferenza di secondo livello. Molte persone non sono affatto sicure delle preferenze che potrebbero avere in un contesto particolare e singolare come quello in cui sono incapacitate; tuttavia, riescono con grande sicurezza a individuare un fiduciario a cui affidarsi non tanto perché il fiduciario sia colui che conosce meglio di altri le preferenze di primo grado del paziente, ma perché la decisione che prenderà – qualunque essa sia – sarà appropriata proprio in quanto presa dal dal fiduciario⁵⁰. Per evitare questa *impasse* potrebbe essere necessario non soltanto nominare il fiduciario, ma esplicitare anche le *ragioni* per le quali è stato scelto il fiduciario: se si farà appello solo all'autorità

⁴⁸ EARP et al. 2024.

⁴⁹ Questo aspetto è stato criticato, tra gli altri, da BALDINI 2018.

⁵⁰ Non discuto i problemi etici che questa prospettiva solleverebbe in relazione agli eccessivi oneri decisionali del fiduciario.

epistemica, allora l'algoritmo predittivo potrebbe – a fronte di informazioni sufficienti – avere la priorità *anche* sul fiduciario esplicitamente nominato; in caso contrario – e a prescindere dai risultati che restituisce l'algoritmo predittivo – dovrà essere il fiduciario a farsi carico della scelta.

In secondo luogo, e in termini più generali, si potrebbe affermare che vada rifiutata la prevalenza algoritmica forte perché l'algoritmo predittivo personalizzato – contrariamente a quanto affermato dai suoi sostenitori – non sarebbe affatto uno strumento per la promozione dell'autonomia del paziente, a differenza di una esplicita nomina di un fiduciario o di una redazione delle DAT. Una recente versione di questa critica è offerta da Annoni⁵¹. In questa prospettiva, la questione dell'autonomia è rilevante non solo per gli algoritmi di tipo predittivo basati su dati sociodemografici, ma anche per quelli di tipo personalizzato. Il sotto-sotto-paragrafo successivo sarà dedicato alla trattazione di questo argomento, data la sua rilevanza per valutare l'appropriatezza della posizione della prevalenza algoritmica forte.

4.2.1. *Algoritmo predittivo delle preferenze personalizzato: una (nuova) questione di autonomia?*

Secondo Annoni, è sostanzialmente dubbio che la decisione sostitutiva in quanto tale – che questa sia presa dal familiare o dall'algoritmo – possa essere uno strumento finalizzato al rispetto e alla promozione dell'autonomia del paziente. Nel suo argomento, Annoni richiama la definizione di autonomia proposta da Beauchamp e Childress, secondo cui per parlare di “decisione autonoma” è necessario soddisfare due condizioni: a) essere in possesso della capacità di agire, comprendendo le informazioni rilevanti e le implicazioni della propria scelta (*agency*); b) essere indipendente da influenze di controllo (*liberty*)⁵². Un paziente incapacitato mancherebbe della nozione di *agency* poiché non disporrebbe della capacità di agire. Annoni ammette che è possibile intendere la nozione di *agency* in termini diacronici e quindi affermare che chi redige le DAT abbia – in un dato contesto e in pieno possesso delle sue capacità – espresso un giudizio autonomo per quelli che Dworkin definisce interessi critici, i quali incarnano i valori, le scelte e le convinzioni che guidano la nostra vita e come noi percepiamo il valore della nostra vita⁵³.

A questo punto, i sostenitori dell'algoritmo predittivo potrebbero osservare che nell'arco della nostra vita ci siamo fatti portatori di valori, di credenze e di visioni comprensive del bene. Le nostre scelte vengono definite autonome proprio perché guidate da tali interessi. Di conseguenza, rispettare l'autonomia di una persona incapace significherebbe prendere delle decisioni che sarebbero state approvate dall'individuo se fosse stato per lui possibile prendere una decisione autonoma. Questa prospettiva è denominata *endorsed life view*⁵⁴.

Annoni rifiuta l'*endorsed life view* poiché, accettando la visione dell'autonomia sottesa a questa prospettiva, questa si dovrebbe applicare anche ad altri contesti decisionali, non solo a quelli riguardanti i pazienti incapaci; ciò condurrebbe a derive di paternalismo medico. In questo senso, per valutare quali trattamenti risultano proporzionati e accettabili per un paziente, il medico potrebbe semplicemente verificare se essi siano in linea con i valori di vita del paziente, anziché basarsi sul consenso informato. Annoni conclude che l'algoritmo predittivo non può essere considerato uno strumento per promuovere l'autonomia, ma piuttosto un mezzo per favorire altri valori, come la coerenza con il piano di vita dell'individuo e l'onorare la sua identità, come nel caso della scelta della canzone che il paziente avrebbe preferito per il proprio funerale: questi sono aspetti importanti ma diversi dalla promozione dell'autonomia. L'autonomia è qui intesa come un atto linguistico performativo: rispettare l'autonomia è dunque concettualmente diverso dal rispetto delle preferenze del paziente.

⁵¹ ANNONI 2024.

⁵² BEAUCHAMP, CHILDRESS 2019.

⁵³ DWORKIN 1993.

⁵⁴ PHILLIPS, WENDLER 2015.

L'argomento di Annoni è persuasivo; tuttavia, si possono muovere almeno tre obiezioni che ne smorzano l'efficacia. La prima obiezione è direttamente rilevante per la discussione sulla posizione della prevalenza algoritmica forte. Come le critiche, già discusse sopra, che si basano sul concetto di autonomia, si ritiene che l'argomento di Annoni funzioni in modo efficace principalmente per alcune proposte di algoritmo predittivo ma non per altre. Se con il PPP è certo che la preferenza predetta non è il risultato di una preferenza effettivamente espressa, con il P4 le cose possono essere diverse. Un commento o un post in un blog potrebbero essere intesi come delle dimostrazioni – talvolta addirittura pubbliche – di una certa volontà del paziente. Riconoscere questo aspetto permetterebbe di conservare una certa dimensione performativa dell'autonomia supportata, anche nell'applicazione dell'algoritmo predittivo personalizzato. Ciononostante, si riconosce che la redazione di una DAT e un *like* sui social network non possono avere lo stesso grado di performatività. In questo contesto, è utile richiamare la differenza proposta da Buchanan e Brock tra una DAT e una preferenza espressa. La prima è un atto di volontà che ha in sé una dimostrazione di impegno che cristallizza un interesse critico, mentre la seconda è semplicemente un atteggiamento propositivo, un desiderio di avere una cosa piuttosto che un'altra, dalla quale non sempre segue un atto di volontà. Gli autori notano che «a volte una persona sceglie deliberatamente di non fare ciò che preferisce, sia per prudenza sia perché ritiene che sarebbe moralmente sbagliato farlo»⁵⁵. Buchanan e Brock ammettono, tuttavia, che questa distinzione è tanto più plausibile quanto più le preferenze sono indirette e inferenziali. A fronte di preferenze chiare ed esplicite grazie a questionari sul tema o dichiarazioni nei blog, social network o siti web, soprattutto se reiterati nel tempo, la distinzione si farebbe meno marcata e quindi meno rilevante.

Queste considerazioni sottolineano la rilevanza della distinzione tra preferenze e scelte autonome, la quale indebolisce la posizione della prevalenza algoritmica forte. In questa prospettiva, al fine di rendere possibile l'identificazione di quelle espressioni di preferenze che soddisfano i requisiti per essere considerate veri e propri "atti di volontà", diventa ancor più rilevante avere accesso in modo diretto alle fonti da cui il predittore algoritmico ricava le proprie risposte. Ciò sottolinea la necessità di promuovere la *spiegabilità* e la trasparenza dell'algoritmo predittivo ai suoi fruitori. Solo in queste situazioni vi sarebbe la possibilità che alcune DAT – come quelle redatte molto tempo prima o quelle che dimostrano delle lacune – siano meno rilevanti rispetto a quanto indicato dall'algoritmo predittivo personalizzato: in tali contesti, non vi sarà un conflitto tra una preferenza e una scelta autonoma, ma tra due "atti di volontà". Tuttavia, si riconosce che questi siano casi sostanzialmente rari: è improbabile che una persona che in precedenza abbia redatto delle DAT e successivamente cambi idea, anche manifestandolo pubblicamente, non provveda a modificarle. Ciò che qui interessa è la possibilità *concettuale* di un conflitto tra due atti di volontà e, di conseguenza, l'applicabilità della posizione della prevalenza algoritmica forte. Alla luce di ciò, sebbene, nei fatti, la critica di Annoni risulti plausibile, da un punto di vista concettuale essa non riesce a confutare la posizione della prevalenza algoritmica forte, poiché è possibile rintracciare atti di volontà almeno in alcune espressioni di preferenze registrate dall'algoritmo predittivo.

La seconda obiezione evidenzia che, nel tentativo di preservare la superiorità delle DAT rispetto al giudizio sostitutivo dell'algoritmo, l'argomento di Annoni apre la strada a scenari problematici nei casi in cui le DAT non siano state redatte. Infatti, se l'utilizzo dell'algoritmo predittivo viene considerato solo come uno strumento finalizzato a onorare l'identità dell'incapacitato e non a promuovere l'autonomia del paziente, allora altri valori potrebbero essere controbilanciati a questi, con la conseguenza controintuitiva che le preferenze del paziente potrebbero risultare non più così centrali; certamente non così importanti come lo erano quando il rispetto delle preferenze del

⁵⁵ BUCHANAN, BROCK 1990, 116.

paziente era inteso come un'espressione dell'autonomia del paziente. In questa prospettiva, se nel giudizio sostitutivo non è possibile rispettare l'autonomia del paziente, con o senza algoritmo, le preferenze *dirette* dei familiari – e. g., quello che i familiari ritengono buono per il paziente – potrebbero assumere un peso molto più importante. Ma è possibile spingersi oltre: onorare l'identità potrebbe essere considerata addirittura meno importante degli *interessi* dei familiari o della struttura sanitaria. In questo senso, verrebbe meno la necessità di rispettare il criterio della fedeltà. In altre parole, la proposta di Annoni nasconde il rischio di depotenziare in modo significativo la richiesta normativa di agire secondo le preferenze che avrebbe avuto il paziente se fosse stato in grado di intendere e di volere. Come già affermato sopra, alcune preferenze possono essere talmente chiare ed esplicite da essere considerate dei veri e propri atti di volontà. Ma anche nel caso in cui non si riuscisse a raggiungere questo standard, ma comunque si riuscisse a delineare in modo sufficientemente adeguato atteggiamenti propositivi, desideri, nonché i valori espressi nella storia biografica del paziente, sembra ragionevole che tali considerazioni debbano essere intese come delle istanze di autodeterminazione. Tali istanze non saranno così forti come quelle contenute negli atti di volontà, ma saranno comunque prioritarie rispetto ad altri interessi in gioco. Da ciò sembra ragionevole preservare all'interno del giudizio sostitutivo una dimensione di "autonomia debole"⁵⁶, quando è possibile rintracciare in modo adeguato le preferenze e dei valori espressi nella storia biografica del paziente.

La terza obiezione rileva che è dubbio che preservare una dimensione di promozione dell'autonomia all'interno del giudizio sostitutivo implichi necessariamente una legittimazione del paternalismo medico nei confronti di pazienti capaci di intendere e di volere. Innanzitutto, vi è una ragione di carattere pratico. In condizioni ordinarie di pratica clinica, il medico non è nella posizione epistemica di conoscere i valori dell'individuo e la coerenza delle sue scelte con il suo piano di vita. In altri termini, il medico non può conoscere le preferenze dell'individuo senza chiedergliele e quindi non può agire in modo informato in base a queste. L'autonomia del paziente sarebbe quindi rispettata solo chiedendo al paziente stesso; solo così, il medico verrebbe a conoscenza delle preferenze del paziente, cristallizzate in un atto di volontà: il consenso informato. Si assuma tuttavia, per amore di discussione, che il medico sia nella posizione epistemica di conoscere le volontà del paziente. Secondo la prospettiva avversata da Annoni, in questo contesto il consenso rischierebbe di essere superfluo, poiché il medico conoscerebbe già le preferenze del paziente e questo lo legittimerebbe a scavalcare la richiesta diretta al paziente. Ma così facendo, il paziente verrebbe privato della facoltà di *rivedere* quelle preferenze, azione che è ancora possibile in quanto il paziente è in grado di intendere e di volere. In caso contrario, si entrerebbe in un cortocircuito per il quale al fine di rispettare i valori e le preferenze del paziente non si permetterebbe al paziente di rivederli. Alla luce di ciò, anche conoscendo le preferenze pregresse del paziente, il medico non potrebbe evitare di chiedere il consenso informato, che rimane un elemento imprescindibile nel rapporto medico-paziente.

Alla luce delle obiezioni qui avanzate, si sostiene che l'algoritmo predittivo può essere considerato uno strumento per la promozione dell'autonomia senza incorrere nelle conseguenze negative prospettate da Annoni. Lo sarà in senso forte quando sarà possibile individuare la dimensione performativa dell'autonomia, attraverso l'individuazione di manifestazioni di volontà reiterate nel tempo, informate ed esplicite. Lo sarà invece in senso debole quando sarà possibile identificare delle preferenze che – seppur non soddisfacendo gli standard più stringenti dell'autonomia forte – sono atteggiamenti e desideri in linea con i valori e la storia personale del paziente.

In questa prospettiva, l'argomento di Annoni non riesce a minare alle sue fondamenta la posizione della prevalenza algoritmica forte, anche se ne limita sensibilmente l'ambito di applicazione. Infatti, nel caso in cui un algoritmo predittivo non fosse allenato su "atti di volontà" –

⁵⁶ BEAUCHAMP, CHILDRESS 2019.

anche se riflettesse in modo adeguato ciò che il paziente avrebbe deciso se avesse potuto esprimersi – non potrebbe prevalere sulle DAT, qualora presenti. Tuttavia, quando l'algoritmo fosse allenato su “atti di volontà”, allora potrebbero esserci circostanze in cui la decisione algoritmica potrebbe prevalere sulle DAT. Questa conclusione, tuttavia, non è da considerarsi definitiva; altri argomenti, sia di natura bioetica sia giuridica, possono essere mossi contro la posizione della prevalenza algoritmica forte, e quindi sono necessarie ulteriori ricerche per valutare il ruolo e il peso della predizione algoritmica nelle decisioni sostitutive, qualora si fosse ragionevolmente certi che l'algoritmo fosse in grado di individuare in modo affidabile le preferenze che il paziente avrebbe avuto nel caso avesse potuto esprimersi.

5. Conclusioni: la necessità del ruolo attivo del paziente e l'IA come supporto a una efficace redazione delle DAT

In questo contributo sono state analizzate alcune delle più rilevanti questioni bioetiche dell'eventuale applicazione dell'algoritmo predittivo personalizzato nelle decisioni sostitutive. In linea di principio, questo strumento ha il potenziale di essere più efficace nel predire le preferenze del paziente incapace rispetto ai familiari e, in alcuni casi, anche rispetto al fiduciario e ad alcune DAT. Tuttavia, l'algoritmo predittivo incontra diversi problemi, come le difficoltà tecniche dovute alla scarsità di dati rilevanti per formulare previsioni attendibili o il rischio di rispecchiare preferenze rivelate ma non effettive e, in alcuni casi, contraddittorie. Inoltre, anche ammettendo una certa affidabilità dell'algoritmo, rimane aperta la questione relativa a quale rilevanza attribuire alla previsione algoritmica nella decisione sostitutiva, quando questa confligge con i giudizi dei familiari, del fiduciario o persino con le DAT.

Per superare alcuni di questi problemi, è stato suggerito di coinvolgere attivamente il paziente, valutando e rivedendo quali informazioni far considerare all'algoritmo, controllando le previsioni e richiedendo il consenso per il trattamento dei dati⁵⁷. Questa posizione è stata abbracciata anche dai proponenti del P4⁵⁸. In altre parole, l'algoritmo predittivo personalizzato richiederebbe al paziente uno sforzo di ragionare in relazione ai suoi valori e alle sue preferenze in condizioni di incapacità.

Tuttavia, se tale sforzo risulta imprescindibile per una implementazione appropriata dell'algoritmo predittivo personalizzato, ci si può legittimamente chiedere se questo non possa essere meglio impiegato per una redazione di DAT più informate ed efficaci. Piuttosto che chiedere al paziente di rivedere dati o possibili preferenze generate dall'algoritmo predittivo personalizzato, potrebbe essere più efficace sviluppare uno strumento che chieda esplicitamente al paziente di esprimere le proprie volontà⁵⁹. Le opportunità offerte dallo sviluppo tecnologico dovrebbero quindi accompagnare il paziente in un nuovo modo di redigere le DAT, anziché limitarsi all'allenamento di un algoritmo predittivo personalizzato.

In questo contesto, si propone lo sviluppo di un nuovo strumento – le DAT dinamiche supportate dall'IA – che, sotto forma di applicazione scaricabile anche sui dispositivi elettronici, possa offrire un supporto per l'identificazione e la conservazione delle disposizioni del paziente. Questo strumento potrebbe svolgere diversi compiti: a) fornire informazioni sulle condizioni di

⁵⁷ In MILIAN, BHATTACHARYYA 2025 è stato proposto un meccanismo che permette agli individui in grado di intendere e di volere di consentire formalmente all'utilizzo di P4.

⁵⁸ EARP et al. 2024.

⁵⁹ In RZEPIŃSKI et al. 2024 viene posta una domanda simile, prevedendo in questo contesto un ruolo centrale per il medico nella raccolta delle DAT. Siccome l'ordinamento italiano non prevede necessariamente il coinvolgimento del medico nella redazione, in questa sede non lo considera necessario, anche se lo si ritiene auspicabile.

incapacità e sui nuovi progressi della medicina in questo ambito⁶⁰; b) attraverso domande mirate, chiedere al paziente le sue preferenze presentando diversi scenari; c) notare eventuali incoerenze nelle preferenze registrate; d) chiedere al paziente di confermare o rivedere le informazioni; e) conservare eventuali dichiarazioni del paziente. Grazie ai dati raccolti sarebbe quindi possibile creare un assistente virtuale spiegabile che potrà essere interrogato dai clinici e dai familiari riguardo alle volontà del paziente in caso di incapacità. In generale, questo strumento permetterebbe di instaurare con il paziente un dialogo nel tempo su queste tematiche, permettendo non solo di preservare uno spazio di autonomia nel caso di incapacità, ma addirittura di potenziarlo rispetto a quanto permettono gli strumenti attualmente disponibili. Al giorno d'oggi, strumenti simili a questo vengono regolarmente utilizzati per il monitoraggio della salute⁶¹ o dei progressi lavorativi; ed è pertanto verosimile prevedere un'applicazione anche per la raccolta delle DAT. Inoltre, attraverso un meccanismo di autenticazione digitale, come il Sistema Pubblico di Identità Digitale, sarebbe possibile conservare le nuove DAT in modo sicuro senza la necessità di conformarsi a quanto affermato dal comma 6 dell'art. 4.

Si ritiene che le DAT dinamiche supportate dall'IA offrano diversi vantaggi sia rispetto al predittore algoritmico personalizzato sia alle DAT ordinarie. La quantità di dati rilevanti ottenuti grazie a questo strumento sarà potenzialmente sufficiente per formulare delle decisioni affidabili in linea con le volontà del paziente, le quali risulteranno aggiornate e informate in misura superiore non solo rispetto a quelle del predittore algoritmico personalizzato, ma anche rispetto ad alcune DAT ordinarie. In questo modo, verrebbero meno anche le preoccupazioni relative all'autonomia discusse nel paragrafo precedente.

Tuttavia, questi vantaggi non devono impedire di osservare che lo strumento proposto può presentare alcuni problemi che dovranno essere adeguatamente indagati attraverso ricerche future per una valutazione esaustiva di una sua possibile introduzione. Ad esempio, sarà necessaria una approfondita valutazione degli eventuali problemi tecnici dello strumento: questi sono in parte sovrapponibili a quelli già evidenziati in precedenza riguardo all'utilizzo dell'IA nell'ambito del predittore algoritmico, i quali assumono particolare rilevanza alla luce della possibilità di interazione tra l'utente e l'IA su questioni delicate e complesse come quelle in esame. In questa prospettiva la possibilità di allucinazioni dell'IA e il rischio che le informazioni offerte varino in base al *prompting* utilizzato per interagirvi potrebbero portare alla trasmissione di indicazioni inesatte o fuorvianti, influenzando il paziente e inducendolo ad avere alcune preferenze rispetto ad altre. Ciò potrebbe compromettere, in ultima istanza, l'autonomia del paziente anziché promuoverla.

Infine, è importante sottolineare che l'introduzione delle DAT dinamiche non eliminerebbe la necessità di decisioni sostitutive, soprattutto nei casi in cui il paziente non abbia utilizzato l'applicazione o espresso esplicitamente delle volontà, anche se potrebbe ridurre sensibilmente il numero. In queste situazioni, strumenti algoritmici predittivi potrebbero comunque rappresentare un utile supporto decisionale per i familiari e i clinici, pur essendo meno efficaci rispetto alle DAT dinamiche supportate dall'IA.

⁶⁰ Si noti che uno dei problemi delle DAT è proprio quello di considerare preferenze che non tengano conto dei progressi medici compiuti nel periodo successivo alla redazione del documento (BUCHANAN, BROCK 1990).

⁶¹ In questa sede, si pensi all'applicazione TrecMamma, un assistente virtuale di TrentinoSalute4.0 rivolta ai neo genitori, per accompagnare e supportare la donna e la famiglia durante la gravidanza e, in futuro, nei primi 1000 giorni di vita del bambino.

Riferimenti bibliografici

- ANNONI M. 2024. *It Is not about Autonomy: Realigning the Ethical Debate on Substitute Judgement and AI Preference Predictors in Healthcare*, in «Journal of Medical Ethics», pubblicato online il 27 novembre 2024. Disponibile in: <https://jme.bmj.com/content/medethics/early/2024/11/27/jme-2024-110343.full.pdf> (consultato il 14/04/25).
- ANYANWU E.C., OKONGWU C.C., OLORUNSOGO T.O., AYO-FARAI O., OSASONA F., DARAOJIMBA O.D. 2024. *Artificial Intelligence in Healthcare: A Review of Ethical Dilemmas and Practical Applications*, in «International Medical Science Research Journal», 4, 2024, 126 ss.
- BALDINI G. 2018. *Prime riflessioni a margine della legge n. 219/17*, in «BioLaw Journal. Rivista di BioDiritto», 2, 2018, 97 ss.
- BALDINI G. 2019. *La legge 219/17 tra molte luci e qualche ombra*, in «Diritti fondamentali», 1, 2019.
- BORSELLINO P. 2018a. *Bioetica tra “moralì” e diritto*, Raffaello Cortina (nuova ed. aggiornata).
- BORSELLINO P. 2018b. *Consenso informato e autodeterminazione terapeutica nella legge sul testamento biologico*, in «Diritto e salute. Rivista di sanità e responsabilità medica», 4, 2018, 35 ss.
- BORSELLINO P. 2019. *La L. n. 219/2017 a un anno dall’approvazione ei nuovi scenari aperti dall’ordinanza della Corte costituzionale n. 207/2018*, in «Notizie di Politeia», 35, 2019, 84 ss.
- BEAUCHAMP T.L., CHILDRESS J.F. 2019. *Principles of Biomedical Ethics*, Oxford University Press (8th ed.).
- BLUMENTHAL-BARBY J., FLETCHER F.E., TAYLOR L., NELSON R.H., MOORE B., SALONER B., UBEL P.A. 2024. *Ethical Complexities in Utilizing Artificial Intelligence for Surrogate Decision Making*, in «The American Journal of Bioethics», 24, 2024, 1 ss.
- BRUDNEY D. 2009. *Choosing for Another: Beyond Autonomy and Best Interests*, in «The Hastings Center Report», 39, 2009, 31 ss.
- BUCHANAN A.E., BROCK D.W. 1990. *Studies in Philosophy and Health Policy: Deciding for Others: The Ethics of Surrogate Decision Making*, Cambridge University Press.
- CAPRON A.M. 2010. *Advance Directives*, in Kuhse H., Singer P. (eds.), *A Companion to Bioethics*, Wiley-Blackwell, 299 ss.
- CHRISTIAN H., SUHARTONO D., CHOWANDA A., ZAMLI K.Z. 2021. *Text Based Personality Prediction from Multiple Social Media Data Sources Using Pre-Trained Language Model and Model Averaging*, in «Journal of Big Data», 8, 2021.
- CIROLDI M., CARIOU A., ADRIE C., ANNANE D., CASTELAIN V., COHEN Y., DELAHAYE A., et al. 2007. *Ability of Family Members to Predict Patient’s Consent to Critical Care Research*, in «Intensive Care Medicine», 33, 2007, 807 ss.
- CNB (COMITATO NAZIONALE PER LA BIOETICA). 2020. *Intelligenza Artificiale e Medicina: Aspetti Etici*.
- CROCKER M., HUTCHINSON C., MPUNDU-KAAMBWA C., WALKER R., CHEN G., RATCLIFFE J. 2021. *Assessing the Relative Importance of Key Quality of Life Dimensions for People with and without a Disability: An Empirical Ranking Comparison Study*, in «Health and Quality of Life Outcomes», 19, 2021, 264 ss.
- D’AMICO M., PELLIZZONE I., LIBERALI B. 2019. *Fase finale della vita, tra libertà ed esigenze di protezione. Quali punti fermi?*, in «Notizie di Politeia», 35, 2019, 71 ss.
- DALTON-BROWN S. 2020. *The Ethics of Medical AI and the Physician-Patient Relationship*, in «Cambridge Quarterly of Healthcare Ethics», 29, 2020, 115 ss.
- DE MARZO G. 2018. *Prime note sulla legge in materia di consenso informato e disposizioni anticipate di*

- trattamento, in «Federalismi.it. Osservatorio di diritto sanitario», 13 marzo 2018.
- DI COSTANZO C. 2019. *La tutela del diritto alla salute del minore. Riflessioni a margine della legge n. 219/2017*, in «BioLaw Journal. Rivista di BioDiritto», 1, 2019, 299 ss.
- DWORKIN R. 1993. *Life's Dominion: An Argument about Abortion and Euthanasia*, Alfred A. Knopf.
- EARP B.D., PORS DAM MANN S., ALLEN J., SALLOCH S., SUREN V., JONGSMA K., BRAUN M., et al. 2024. *A Personalized Patient Preference Predictor for Substituted Judgments in Healthcare: Technically Feasible and Ethically Desirable*, in «The American Journal of Bioethics», 24, 2024, 13 ss.
- FERRANDO G., BUSATTA L., FRANCESCONI A., BERGONZINI C., MOLASCHI V., VERONESI P., BENCIOLINI P., GIARDINA F., PIZZETTI F.G., CANESTRARI S., MORINO P., VIAFORA C., GIANNINI A., ZAMPERETTI N., PARIS D., GRISTINA G.R., ORSI L. 2018. *La legge n. 219 del 2017, Norme in materia di consenso informato e di disposizioni anticipate di trattamento*, in «BioLaw Journal. Rivista di BioDiritto», 1, 2018, 19 ss.
- FERRARIO A., GLOECKLER S., BILLER-ANDORNO N. 2023. *Ethics of the Algorithmic Prediction of Goal of Care Preferences: From Theory to Practice*, in «Journal of Medical Ethics», 49, 2023, 165 ss.
- FLORIDI L. 2023. *The Ethics of Artificial Intelligence*, Oxford University Press.
- GIOVANOLA B., TIRIBELLI S. 2023. *Beyond Bias and Discrimination: Redefining the AI Ethics Principle of Fairness in Healthcare Machine-Learning Algorithms*, in «AI & Society», 38, 2023, 549 ss.
- HOUTS R.M., SMUCKER W.D., JACOBSON J.A., DITTO P.H., DANKS J.H. 2002. *Predicting Elderly Outpatients' Life-Sustaining Treatment Preferences over Time: The Majority Rules*, in «Medical Decision Making», 22, 2002, 39 ss.
- HUBBARD R., GREENBLUM J. 2020. *Surrogates and Artificial Intelligence: Why AI Trumps Family*, in «Science and Engineering Ethics», 26, 2020, 3217 ss.
- JOHN S., ROWLEY J., BARTLETT K. 2020. *Assessing Patients Decision-Making Capacity in the Hospital Setting: A Literature Review*, in «The Australian Journal of Rural Health», 28, 2020, 141 ss.
- KARIMIAN G., PETELOS E., EVERS S.M.A.A. 2022. *The Ethical Issues of the Application of Artificial Intelligence in Healthcare: A Systematic Scoping Review*, in «AI and Ethics», 2, 2022, 539 ss.
- KOSINSKI M., STILLWELL D., GRAEPEL T. 2013. *Private Traits and Attributes Are Predictable from Digital Records of Human Behavior*, in «Proceedings of the National Academy of Sciences of the United States of America», 110, 2013, 5802 ss.
- LAMANNA C., BYRNE L. 2018. *Should Artificial Intelligence Augment Medical Decision Making? The Case for an Autonomy Algorithm*, in «AMA Journal of Ethics», 20, 2018, E902 ss.
- LENZI D. 2018. *La legge 219/2017 e il difficile percorso parlamentare*, in «Diritto e salute. Rivista di sanità e responsabilità medica», 2018, 99 ss.
- LEVINE D.M., TUWANI R., KOMPA B., VARMA A., FINLAYSON S.G., MEHROTRA A., BEAM A. 2024. *The Diagnostic and Triage Accuracy of the GPT-3 Artificial Intelligence Model: an Observational Study*, in «The Lancet Digital Health», 8, 2024, e555-e561.
- LORENZINI G., ARBELAEZ OSSA L., SHAW D.M., ELGER B.S. 2023. *Artificial Intelligence and the Doctor-Patient Relationship Expanding the Paradigm of Shared Decision Making*, in «Bioethics», 37, 2023, 424 ss.
- MANNELLI C. 2022. *Etica e Intelligenza artificiale. Il caso sanitario*, Donzelli.
- MILIAN, R. D., & BHATTACHARYYA, A. 2025. *Beyond Presumed Autonomy: AI-Assisted Patient Preference Predictors and the Personalised Living Will*, in «Journal of Medical Ethics», 2025. Disponibile in: <https://doi.org/10.1136/jme-2025-110718> (consultato il 14/05/25).
- MORLEY J., MACHADO C.C.V., BURR C., COWLS J., JOSHI I., TADDEO M., FLORIDI L. 2020. *The Ethics of AI in Health Care: A Mapping Review*, in «Social Science & Medicine», 260, 2020, 113172.

- MURPHY K., DI RUGGIERO E., UPSHUR R., WILLISON D.J., MALHOTRA N., CAI J.C., MALHOTRA N., LUI V., GIBSON J. 2021. *Artificial Intelligence for Good Health: A Scoping Review of the Ethics Literature*, in «BMC Medical Ethics», 22, 2021, 14.
- NAIK N., HAMEED B.M.Z., SHETTY D.K., SWAIN D., SHAH M., PAUL R., AGGARWAL K., et al. 2022. *Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility?*, in «Frontiers in Surgery», 9, 2022, 862322.
- ONEIL C. 2022. *Commentary on 'Autonomy-Based Criticisms of the Patient Preference Predictor'*, in «Journal of Medical Ethics», 48, 2022, 315 ss.
- OSSERVATORIO PERMANENTE SULLE DAT 2024. *I Testamenti Biologici nei Comuni Italiani*. Disponibile in https://lookerstudio.google.com/u/o/reporting/856a479b-25eb-4850-8dcd-5fd595dic2f8/page/p_typhw3svad (Consultato il 14/04/25).
- PALAZZANI L. 2018. *Le DAT e la legge 219/2017: considerazioni bioetiche e biogiuridiche*, in «Rivista Italiana di Medicina Legale (e del Diritto in campo sanitario)», 3, 2018, 965 ss.
- PALAZZANI L. 2020. *Tecnologie dell'informazione e intelligenza artificiale: Sfide etiche al diritto*. Edizioni Studium.
- PHILLIPS J., WENDLER D. 2015. *Clarifying Substituted Judgement: The Endorsed Life Approach*, in «Journal of Medical Ethics», 41, 2015, 723 ss.
- REICHLIN M. 2021. *Fondamenti di Bioetica*, il Mulino.
- RID A., WENDLER D. 2014. *Treatment Decision Making for Incapacitated Patients: Is Development and Use of a Patient Preference Predictor Feasible?*, in «The Journal of Medicine and Philosophy», 39, 2014, 130 ss.
- RZEPIŃSKI T., DESKUR-ŚMIELECKA E., CHOJNICKI M. 2024. *Machine Learning Algorithms in the Personalized Modeling of Incapacitated Patients' Decision Making – Is It a Viable Concept?*, in «The American Journal of Bioethics: AJOB», 24, 2024, 51 ss.
- SAVULESCU J., GIUBILINI A., VANDERSLUIS R., MISHRA A. 2024. *Ethics of Artificial Intelligence in Medicine*, in «Singapore Medical Journal», 65, 2024, 150 ss.
- SHALOWITZ D.I., GARRETT-MAYER E., WENDLER D. 2006. *The Accuracy of Surrogate Decision Makers: A Systematic Review*, in «Archives of Internal Medicine», 166, 2006, 493 ss.
- SHALOWITZ D.I., GARRETT-MAYER E., WENDLER D. 2007. *How Should Treatment Decisions Be Made for Incapacitated Patients, and Why?*, in «PLoS Medicine», 4, 2007, e35.
- SHARADIN N. 2018. *Patient Preference Predictors and the Problem of Naked Statistical Evidence*, in «Journal of Medical Ethics», 44, 2018, 857 ss.
- SHARADIN N. 2024. *Personalized Patient Preference Predictors Are Neither Technically Feasible nor Ethically Desirable*, in «The American Journal of Bioethics: AJOB», 2024.
- SMUCKER W.D., HOUTS R.M., DANKS J.H., DITTO P.H., FAGERLIN A., COPPOLA K.M. 2000. *Modal Preferences Predict Elderly Patients' Life-Sustaining Treatment Choices as Well as Patients' Chosen Surrogates Do*, in «Medical Decision Making», 20, 2000, 271 ss.
- YADAV K.N., GABLER N.B., COONEY E., KENT S., KIM J., HERBST N., MANTE A., HALPERN S.D., COURTRIGHT K.R. 2017. *Approximately One in Three US Adults Completes Any Type of Advance Directive for End-of-Life Care*, in «Health Affairs», 36, 2017, 1244 ss.
- YOUYOU W., KOSINSKI M., STILLWELL D. 2015. *Computer-based personality judgments are more accurate than those made by humans*, in «Proceedings of the National Academy of Sciences», 112, 2015, 1036 ss.