

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

**Dottorato di ricerca in Scienze Mediche Generali e
Scienze dei Servizi**

Ciclo XXX

Settore Concorsuale: 06/M1

Settore Scientifico Disciplinare: MED/42

TITOLO TESI

Il ruolo degli indicatori patient reported (PROMs) per la promozione della qualità dell'assistenza: il caso di studio della chirurgia del CA prostatico.

Presentata da: Dott. Davide Gori

Coordinatore Dottorato

Prof. Fabio Piscaglia

Supervisore

Prof.ssa Maria Pia Fantini

Esame finale anno 2018

Abstract della tesi in lingua italiana:

La definizione della qualità dell'assistenza sanitaria richiede una conoscenza e un approfondimento di quanto le persone beneficino dei servizi sanitari, richiede di misurare fattori quali gli esiti di specifiche procedure diagnostiche e terapeutiche, e di misurare la soddisfazione del paziente e più generalmente il senso di benessere della società.

Poiché è estremamente difficile mantenere la qualità in un contesto di riduzione di costi e più in generale delle risorse, ognuno di questi aspetti riceve uguale attenzione negli attuali dibattiti scientifici, in particolare nel contesto dei moderni sistemi sanitari integrati. Le sempre maggiori sfide per il futuro saranno quindi quelle di contenere i costi e contemporaneamente migliorare la qualità.

A partire dalla seconda metà degli anni 2000, il servizio Sanitario Inglese e i servizi sanitari di molti paesi in area OCSE, hanno iniziato, sia per scopi di valutazione della qualità delle cure erogate, sia per ragioni di *accountability* e di valutazione della performance dei sistemi sanitari, a raccogliere indicatori riportati dai pazienti. Questi indicatori possono essere grossolanamente divisi in indicatori riguardanti gli esiti riferiti dai pazienti, sia indicatori riguardanti la loro esperienza di cura. In dettaglio esse sono le misure di esperienza riportate dal paziente (PREMs – Patient Reported Experience Measure, come ad esempio se il paziente ritenga di essere adeguatamente coinvolto in decisioni importanti circa il suo processo di cura) sia alle misure di esito del paziente (PROM – Patient Reported Outcome Measures, ad esempio se il paziente è libero dal dolore/ha un livello di dolore sopportabile dopo una operazione chirurgica).

I PROMs, in particolare, hanno visto negli ultimi anni un aumento esponenziale del loro utilizzo a tutti i livelli dell'assistenza, e in particolare per l'assistenza alle patologie croniche con particolare riferimento alle patologie oncologiche. Molti studi recenti suggeriscono comunque la necessità di integrare la misurazione delle prestazioni in un programma formale e più generale di miglioramento della qualità.

Il cancro della prostata è la diagnosi più comune di tumore maligno e la seconda causa di morte negli uomini. Gli ultimi 20 anni hanno visto un calo di oltre il 50% dei tassi di mortalità per cancro alla prostata in virtù di una migliorata diagnosi e trattamento precoce. Anche se il controllo sulla progressione del cancro e sui tassi di sopravvivenza dopo la prostatectomia sono eccellenti, i pazienti possono sviluppare effetti collaterali legati al trattamento.

Le metriche di qualità per gli effetti collaterali vanno ad indirizzarsi specialmente verso l'analisi di eventuali differenze di trattamento o su esiti diversi e includono misure di struttura, di processo e di esito per l'assistenza che viene erogata.

Nella prima parte sperimentale della presente tesi sono stati identificati gli indicatori più importanti di qualità per il cancro alla prostata (proposti dalla comunità scientifica) e determinarne il loro utilizzo.

Il nostro studio oltre ad enucleare i 10 indicatori più “comunemente misurati” ha individuato 5 indicatori fondamentali (i più utilizzati e impattanti) nell'ambito della chirurgia oncologica per il carcinoma della prostata:

- *“Valutazione del funzionamento urinario, sessuale e intestinale del paziente dopo il trattamento”;*
- *“Soddisfazione dei pazienti per la scelta del trattamento, per la continenza e la potenza sessuale”*

Il primo di essi è stato misurato e confrontato con il reporting dei clinici in un caso di studio concreto di un ospedale universitario americano – Stanford, California.

Il mancato sviluppo e utilizzo di misure di outcome può rappresentare quindi un limite allo sviluppo di futuri programmi nazionali o regionali, e al coinvolgimento dei pazienti con la valutazione di esiti davvero impattanti nella loro esperienza di malattia e nella relativa qualità di vita percepita.

L'adozione diffusa di cartelle cliniche elettroniche (EHR) offre l'opportunità di estrarre informazioni sugli esiti correlati al trattamento dalle note cliniche regolarmente registrate come testo libero nella EHR. L'avvento di strumenti sempre più sofisticati, come

l'elaborazione del *natural language processing* (NLP), può essere utilizzato per identificare ed estrarre in modo efficiente dati non strutturati dalle EHR su esiti *patient-centered* e riportati dai medici.

Abbiamo sviluppato un flusso di estrazione e di *data-mining* che è stato in grado di catturare la documentazione fornita dai medici sugli esiti più importanti e centrati sul paziente nei pazienti con cancro alla prostata utilizzando note cliniche non strutturate e registrate di routine dalle EHR. Il nostro strumento ha identificato in modo efficiente e accurato le valutazioni dell'incontinenza urinaria dopo la prostatectomia e abbiamo riscontrato che la segnalazione da parte dei medici dei sintomi della IU è altamente concordante con quanto viene riferito dai pazienti utilizzando l'indagine EPIC-26, un importante PROM che nel corso degli anni è stato validato ed è attualmente considerato uno dei *gold standard* per la valutazione del paziente post-prostatectomia.

L'incontinenza urinaria (IU) e la disfunzione erettile (DE) in particolare sono gli effetti collaterali più frequenti della PR e quelli che influenzano maggiormente la qualità della vita post intervento.

Il caso di studio presentato (relativo alla valutazione della IU post intervento per carcinoma prostatico e alla sperimentazione e alla raccolta dati sugli esiti funzionali estratti dalle cartelle cliniche informatizzate utilizzando i dati del database Stanford University HealthCare), apre la strada alla possibilità di utilizzare e valutare questo tipo di indicatori in maniera sistematica traendo informazioni dalla grande quantità di dati a disposizione in formato elettronico nell'ambito dei servizi sanitari. I clinici, infatti, stanti i nostri risultati nell'ambito degli esiti nella patologia oncologica prostatica (ma come da noi detto probabilmente questo dato sarà trasferibile anche in altri contesti) potrebbero riportare nelle cartelle cliniche importanti informazioni riferite dai pazienti. Ciò dovrà essere unito a programmi di formazione e di implementazione della raccolta dei PROMs in modo da convincere i clinici – che sono sicuramente più propensi ad affidarsi a *endpoints* “hard” (quali la morte, la sopravvivenza, etc.) – a valutare e a raccogliere nelle loro cartelle, in modo sistematico, informazioni riferite a questi esiti funzionali.

Questo cambiamento potrà portare, in futuro, ad una migliore raccolta dei dati sugli esiti *patient-centered* e ad un effettivo miglioramento della qualità così orientata in tutti i settori della medicina.

Sommario

La qualità nell'assistenza e la metrica Donabediana.....	9
PREMs e PROMs.....	13
L'uso di misure di esperienza riportate dal paziente (PREMs)	14
I sistemi sanitari stanno applicando sempre più PREMs come indicatori critici delle prestazioni.....	17
Le opportunità di confronto tra PREMs.....	20
L'utilizzo di misure di esito del paziente (PROMs) nei sistemi sanitari.....	22
Tipi di PROMs.....	23
PROM generici relativi alla qualità della vita	23
EQ-5D.....	23
Short Form 36 (SF-36).....	24
SF-12.....	24
Crosswalk algorithms.....	25
Esempi di uso di registri clinici della qualità in alcuni paesi.....	26
L'uso dei PROMS in settori specifici	28
Chirurgia elettiva.....	28
Servizi di salute mentale	28
Degenza a lungo termine.....	29
Cure palliative	30
Informal care	31
Performance measurement statali per diverse condizioni patologiche.....	31
Uso dei PROMS	33
Raccolta e pubblicazione dei dati PROMS	39
La raccolta dei PROMS	42
Impatto dei PROMS.....	44
Caso di studio: indicatori del cancro alla prostata.....	47
Metodi	49

Figura 1: Flow chart che mostra il processo di selezione degli articoli.....	50
Risultati	52
Tabella 1. Caratteristiche descrittive delle misure di qualità per il carcinoma prostatico	54
Tabella 2A. Misure di struttura per il cancro alla prostata utilizzate in letteratura.....	61
Tabella 2B. Misure di processo per il cancro alla prostata utilizzate in letteratura	62
Tabella 2C. Misure di esito per il cancro alla prostata utilizzate in letteratura	64
Tabella 3. “Top-ten” delle misure di qualità per il cancro alla prostata più usate.....	67
Figura 2: Distribuzione stratificata delle misure “top-ten”, a seconda che gli studi fossero focalizzati alla ricerca o al miglioramento clinico.	68
Discussione	70
Tabella 4 Proposta di una serie di misure di qualità essenziali emerse dalla letteratura scientifica.....	73
Caso di studio: potenzialità per la raccolta dei dati PROMs dalle cartelle cliniche	74
Metodi	76
Il questionario EPIC 26	76
Sorgente dei dati.....	76
Popolazione in studio.....	77
Esiti misurati e covariate prese in considerazione per questa analisi	78
Analisi statistica	80
Risultati	80
Figura 3: Flowchart per la selezione della coorte	81
Tabella 5: caratteristiche cliniche e demografiche al baseline	82
Tabella 6: Sensibilità e specificità della definizione di IU ai diversi punti*	85
Figura 4: prevalenza di IU dalle cartelle cliniche e da EPIC-26 al baseline e ai vari punti del follow-up	86
Tabella 7: Score di IU da EPIC-26 nei differenti punti e comparazione con studi condotti in precedenza.	87
Tabella 8: Percentuale di accordo osservato tra EPIC-26 e cartelle cliniche, accordo atteso per effetto casuale da solo e Cohen’s Kappa.	89
Tabella 9. Predittori di disaccordo sulla presenza di IU tra EPIC e EHR. Risultati dal modello di regressione logistica univariata.	90

Tabella 10: Tabella incrociata della IU riportata dai clinici e dai pazienti a 3 mesi, e stratificati per età del paziente.....	91
Discussione.....	92
Conclusioni	97
Bibliografia	100

La qualità nell'assistenza e la metrica Donabedian

La definizione della qualità dell'assistenza sanitaria richiede di conoscere a fondo quanto le persone beneficino dei servizi sanitari, di misurare fattori quali gli esiti di specifiche procedure diagnostiche e terapeutiche, e di misurare la soddisfazione del paziente e il senso di benessere della società. Poiché è estremamente difficile mantenere la qualità in un contesto di riduzione di costi e più in generale delle risorse, ognuno di questi aspetti riceve uguale attenzione negli attuali dibattiti scientifici, in particolare nel contesto dei moderni sistemi sanitari integrati. Le sempre maggiori sfide per il futuro saranno quindi quelle di contenere i costi e contemporaneamente migliorare la qualità.

Nel suo articolo del 1966, ad oggi considerato una pietra miliare su questo argomento, Avedis Donabedian propose di utilizzare una triade di misure per valutare la qualità dell'assistenza sanitaria, classificate secondo il framework concettuale di: "struttura", "processo" ed "esito". (Donabedian, 1966)

Egli definì la "struttura" come il setting in cui avviene il processo di cura, e il modo in cui i providers dei servizi sanitari e i sistemi amministrativi sono qualificati; "processo" come i componenti delle cure erogate; e "esito" come il "recupero", in particolare il recupero della/e funzione/i o, più genericamente, la sopravvivenza. Tutti questi concetti rimangono ancora ad oggi il fondamento della valutazione della qualità dei servizi sanitari e delle cure da essi erogate.

La struttura è stata a lungo valutata attraverso la certificazione delle attività dei medici e l'accreditamento degli ospedali da parte delle agenzie sanitarie pubbliche. Le misure di processo e di esito, negli ultimi anni, sono state esaminate e valutate approfonditamente dalla maggioranza dei panels nazionali per la valutazione della qualità, con un reporting pubblico (per tutte queste misure) a disposizione di ognuno, sulle prestazioni di ospedali, sulle case di cura private e più in generale per la qualità dei piani sanitari. Inoltre, i programmi di certificazione per i medici ora si sforzano di

promuovere l'apprendimento continuo (ECM) e il monitoraggio delle misure di prestazione clinica.

Molto prima dell'avvento delle cartelle cliniche elettroniche o dei dati sanitari computerizzati, l'articolo di Donabedian aveva anche presagito molte delle sfide che persistono ancora oggi e gli attuali sforzi per valutare la qualità delle cure. Donabedian sottolineava infatti le difficoltà di valutare la qualità sulla base di cartelle cliniche o diretta osservazione di incontri clinici. Sottolineava inoltre l'importanza di avere campioni rappresentativi, e standard di misurazione chiari. Nel suo paper sottolineava anche la necessità di possedere valide misure di struttura e processo che potessero, tuttavia, essere collegate direttamente agli esiti. Inoltre affermava come dovessero essere sviluppate misure affidabili di qualità che fossero prontamente riproducibili in diversi contesti. Richiedeva un approccio più ampio alla misurazione della qualità che si estendesse oltre la gestione tecnica della malattia, con l'attivo incorporamento delle valutazioni di prevenzione, riabilitazione, coordinamento, qualità della vita e continuità delle cure. Donabedian poneva particolare accento proprio sulla relazione medico-paziente, sulla efficienza economica e sui valori sociali. E già dai suoi primi scritti prevedeva anche il coinvolgimento attivo dei pazienti (focus primario del processo di cura) nella valutazione della qualità dell'assistenza.

Il focus attuale sul pagamento basato sul valore della cura e sugli esiti centrati sul paziente può essere, quindi, come detto, fatto risalire al quadro concettuale di Donabedian. Questa struttura ha fornito le basi per un rapporto che ha segnato un'epoca, prodotto dall'American Institute of Medicine nel 1990, intitolato "*Medicare: A Strategy for Quality Assurance*". (Lohr, 1991)

Poggiando le sue basi ampiamente sul lavoro di Donabedian, il rapporto definisce la qualità dell'assistenza come "il grado in cui i servizi sanitari per gli individui e le popolazioni aumentano la probabilità di risultati sanitari desiderati e sono coerenti con le

attuali conoscenze professionali". E con ciò ponendo quindi il focus sia sulla qualità (da tutti i punti di vista) sia sulla scientificità e la professionalità che deve essere alla base di tutte le decisioni che vengono prese.

Sempre nel 1990, Donabedian ha delineato i "sette pilastri" della qualità. Quattro di essi vennero individuati 11 anni dopo nel rapporto dell'Institute of Medicine: "*Crossing the Quality Chasm*", che ha evidenziato sei obiettivi fondamentali per il sistema sanitario del XXI secolo, ovvero fornire cure che siano:

- sicure,
- efficaci,
- incentrate sul paziente,
- tempestive,
- efficienti
- eque.

I benchmark relativi a questi obiettivi sono diventati parametri comuni in cliniche, ospedali e organizzazioni di assistenza sanitaria in tutto il mondo.

Donabedian ha concluso il suo articolo del 1966 con una domanda essenziale e una richiesta ai suoi lettori: "Sempre più spesso c'è bisogno di chiedersi: "*Cosa succede qui?*" invece di "*Cosa c'è di sbagliato? E come può essere migliorato?*"

Ciò non significa che il ricercatore rinneghi i propri valori o obiettivi sociali. Significa, tuttavia, che la distinzione tra valori e elementi di struttura, processo o esito è riconosciuta e mantenuta; ma che entrambi sono sottoposti a uno studio altrettanto critico.

Donabedian ha sottolineato che l'attenzione alle metriche è essenziale. Ha anche insistito sul fatto che i medici devono riconoscere che la cultura e i sistemi sociali in cui

praticano possono migliorare o diminuire la qualità dell'assistenza sanitaria. Pertanto, crediamo che Donabedian volesse più che altro invitare i professionisti della salute nel nostro attuale periodo storico a impegnarsi attivamente nella guida e nella formazione della ricerca di un'assistenza di alta qualità nelle loro organizzazioni e rivolta a tutta la società, in primis al paziente, vero centro della cura.

PREMs e PROMs

A partire dalla seconda metà degli anni 2000, il servizio Sanitario Inglese e i servizi sanitari di molti paesi in area OCSE, hanno iniziato, sia per scopi di valutazione della qualità delle cure erogate, sia per ragioni di accountability e di valutazione della performance dei sistemi sanitari, a raccogliere indicatori riportati dai pazienti. Questi indicatori possono essere grossolanamente divisi in indicatori riguardanti gli esiti riferiti dai pazienti, sia indicatori riguardanti la loro esperienza di cura. (Report OECD, 2017)

In dettaglio esse sono le misure di esperienza riportate dal paziente (PREMs – Patient Reported Experience Measure, come ad esempio se il paziente ritenga di essere adeguatamente coinvolto in decisioni importanti circa il suo processo di cura) sia alle misure di esito del paziente (PROM – Patient Reported Outcome Measures, ad esempio se il paziente è libero dal dolore/ha un livello di dolore sopportabile dopo una operazione chirurgica).

Molti dei paesi dell'area OCSE conducono indagini PREMs e, in misura minore, sperimentano i PROMs. L'OCSE ha già molti indicatori comparabili sull'esperienza del paziente, sviluppati nell'ambito della sua raccolta periodica di dati sugli indicatori di qualità. Le statistiche raccolte regolarmente, tuttavia, forniscono un quadro insufficiente degli outcomes dell'assistenza sanitaria. Inoltre, i paesi stanno sovente lavorando in modo indipendente e non trasversale e condiviso, presentando ben poche opportunità per effettuare una analisi comparata tra diversi paesi. In questo capitolo forniremo una sintesi di come gli indicatori riportati dal paziente (sia PREMs che PROMs) possano essere sviluppati e utilizzati e come la raccolta di molte di queste metriche possa in futuro portare vantaggi ai policy makers, ai pazienti e ai medici clinici.

L'uso di misure di esperienza riportate dal paziente (PREMs)

La misura e il controllo dell'esperienza del paziente può indirizzare consapevolmente o suggerire modifiche alla pratica clinica, modifiche necessarie al miglioramento della qualità delle cure erogate. Prendere in considerazione la “voce del paziente” – e questo vale anche per i PROMs - nella progettazione dei sistemi sanitari può anche aiutare nella fornitura di un'assistenza sanitaria che sia più rispondente alle vere esigenze dei pazienti. L'invecchiamento della popolazione e il sempre maggiore carico di malattie croniche e degenerative, unite alla sempre crescente enfasi posta sul self-management del paziente, renderanno ancora più importante la cattura ottimale delle metriche sull'esperienza del paziente in futuro.

I PREMs misurano le percezioni dei pazienti della loro esperienza, concentrandosi sul processo di cura e su come esso sia in grado di influenzare la loro esperienza.

A titolo di esempio possiamo portare quesiti PREMs quali:

“Il paziente ha aspettato molto tempo per il trattamento?” oppure “Il paziente ha ritenuto di essere coinvolto nel processo decisionale?”

Mentre molte attività di PREM si verificano a livello nazionale, esistono tuttavia pochi strumenti che facilitino o si siano occupati di confronti internazionali PREMs. Questa mancanza di informazioni è un'opportunità sprecata, in particolare per i policy makers, di comprendere quelle che sono esperienze insoddisfacenti nell'assistenza in contesti diversi e transnazionali.

Un'esperienza positiva del paziente dovrebbe essere considerata di per sé un outcome di fondamentale importanza, e correlare bene ad altre misure di qualità. C'è evidenza del rapporto tra l'esperienza del paziente con il processo di cura e gli esiti, per cui i PREMs correlano strettamente anche con i PROMs.

Ad esempio, uno studio ha scoperto che - dopo aver aggiustato per lo stato di salute post-dimissione e altri fattori clinici - i pazienti che hanno sperimentato e riferito peggiori cure ospedaliere ed esperienze peggiori, valutavano peggio anche la propria salute

generale, la propria salute fisica e avevano più probabilità di avere dolore toracico un anno dopo un infarto miocardico acuto (IMA) rispetto ai pazienti che fungevano da controllo (migliori PREMs). (Fremont, 2001)

Gli stessi autori sottolineano anche come l'associazione tra un'esperienza ospedaliera negativa e il dolore al torace successivo all'evento IMA possa essere compensata da esperienze ambulatoriali e sul territorio più positive. Sul territorio e fuori quindi dal contesto ospedaliero, i problemi più frequenti si sono verificati nell'informazione e nell'educazione del paziente, nell'assistenza emotiva e psicologica, nel coinvolgimento familiare e degli amici, nella continuità e nel passaggio dai servizi sanitari al proprio domicilio. Gli autori dello studio citano la qualità della comunicazione con i pazienti come fattore chiave. I pazienti che hanno difficoltà ad ottenere una "guida" chiara sulla loro condizione e soprattutto sul trattamento loro prescritto sperimentano una minore probabilità di assumere correttamente i farmaci, modificare lo stile di vita atto a migliorare il loro recupero (ad esempio dieta e attività fisica prescritte) e potrebbero avere meno probabilità di rispettare gli appuntamenti ambulatoriali programmati sul follow-up o di segnalare preoccupazioni/problemi sulla loro salute al medico che li ha in carico sul territorio. Questi pazienti possono anche essere a maggior rischio di ansia o depressione, che sono sicuramente associati a risultati peggiori dopo un IMA (Fremont, 2001).

Uno studio sulle donne con cancro al seno ha mostrato un'associazione tra la terapia tumorale in corso con tamoxifene (quattro anni dopo l'inizio del trattamento) e la misurazione dell'esperienza centrata sul paziente. La percentuale di pazienti che continuavano ad usare tamoxifene era minore per coloro che avevano meno sostegno, un ruolo meno importante nei processi di *decision-making* o pochi input su come prendere decisioni riguardanti l'uso del tamoxifene da parte del medico e infine se non era stato detto loro in anticipo (da uno o più operatori sanitari) quali fossero gli effetti collaterali del farmaco. Questo dato veniva confermato anche dopo aver aggiustato per

la gravità degli effetti collaterali e per tutti gli altri fattori demografici e clinici (Kahn, 2007).

Un altro studio ha trovato una correlazione positiva, seppur modesta, tra misure di esperienza del paziente con misure di processo di qualità clinica nella prevenzione e nella gestione delle malattie nell'assistenza primaria. Tuttavia, non esistevano correlazioni significative tra esperienze di cura del paziente e risultati clinici (Sequist, 2008).

I sistemi sanitari stanno applicando sempre più PREMs come indicatori critici delle prestazioni

Molte iniziative di PREM sono state intraprese nei decenni scorsi, spesso modellate sul lavoro dell'Istituto Picker nel Regno Unito e sulle valutazioni del Consumer Assessment of Healthcare Providers and Systems (CAHPS) negli Stati Uniti. L'Istituto Picker ha sviluppato nel 1987 i principi di Care Centered Care, che è diventato un framework utilizzato a livello internazionale per supportare una cura di alta qualità per i pazienti. Nel 2002, ha progettato e istituito il primo programma nazionale di indagine del National Health Service (NHS) in Inghilterra per valutare l'esperienza del paziente. L'iniziativa CAHPS è stata avviata nel 1995 dalla AHRQ americana, con questionari standardizzati che misurano l'esperienza del paziente.

A livello nazionale, in quasi tutti i paesi dell'OCSE i PREMs sono raccolti attraverso indagini che comprendono campioni di popolazione di pazienti che sperimentano cure ospedaliere o territoriali. Alcuni paesi stanno sviluppando strumenti per valutare l'esperienza di pazienti che hanno specifiche esigenze di cura, in modo da migliorare l'assistenza verso particolari gruppi di popolazione. Queste possono essere indagini specifiche sulla condizione patologica (ad es. cancro o diabete) o specifiche per il percorso di cura (ad esempio per ciò che riguarda l'assistenza materno infantile o quella psichiatrica).

A titolo di esempio, la Norvegia conduce indagini che riguardano persone con malattie specifiche, compresi i pazienti adulti e pediatrici che hanno ricevuto assistenza sanitaria per problemi mentali. Nei Paesi Bassi, i PREMs sono raccolti per persone con diabete, asma, insufficienza cardiaca e cancro, che coprono providers quali medici generici, fisioterapisti, ospedali e case di cura (Fujisawa, 2017). Ci sono molte condizioni per le quali i PREM non vengono raccolti in maniera sistematica. Ad esempio, l'esperienza dei pazienti con artrite reumatoide non viene misurata normalmente nella maggioranza dei

paesi. Recentemente è stato sviluppato, condotto e validato un questionario PREMs per l'artrite reumatoide ed è stato utilizzato in un Audit Nazionale Clinico in Inghilterra e Galles (Bosworth et al., 2015). Non ci sono inoltre PREMs che permettano una valutazione dell'esperienza del paziente con condizioni croniche multiple/multimorbosità.

In Francia l'iniziativa e-Satis misura la soddisfazione del paziente e l'esperienza negli ospedali. È importante sottolineare come questa indagine includa anche domande sul coordinamento dell'assistenza, comprese le domande sulla dimissione ospedaliera e su come viene coordinata la cura tra ospedali e medici di base. Le informazioni vengono riconsegnate agli ospedali per aiutarli a migliorare la qualità. E-Satis fornisce anche informazioni e possibilità di scelta al pubblico. I dati (pubblici e liberamente scaricabili) vengono anche utilizzati allo scopo di pagamento delle prestazioni, in quanto gli ospedali ricevono bonus per risultati di buona qualità. Vi sono dati sia a livello regionale che dati a livello nazionale. I dati di e-Satis sono stati pubblicati per la prima volta nel 2016. Se gli ospedali non ottengono risultati soddisfacenti, viene fatta una nota sul loro registro di accreditamento, che è pubblico. Da notarsi come l'accreditamento è obbligatorio per tutti gli ospedali, sia pubblici che privati, in Francia.

Le misure di esperienza del paziente a livello dei servizi sanitari vengono utilizzate per informare i legislatori sanitari al riguardo delle misure di regolamentazione e accreditamento. Per esempio, la Repubblica Ceca aggiudica i certificati "Satisfied Patient" a quei providers che raggiungono prestazioni eccezionali relative alle esperienze del paziente (Fujisawa, 2017).

Il coordinamento della cura nel percorso del paziente è fondamentale per ridurre il rischio di contrattempi causati da errori di comunicazione tra i differenti providers. Questa è un'area in cui ci sarebbe un grande vantaggio in un maggiore apprendimento sulla base dell'esperienza del paziente, ma in realtà le conoscenze sono ancora piuttosto limitate.

In uno studio statunitense, è stata sviluppata una misura self-reported sulla qualità della cura del paziente nel passaggio tra ospedale e abitazione privata. Questa misura, definita Care Transitions Measure (CTM) ha dimostrato avere un'alta coerenza interna e affidabilità. CTM è stato sviluppato con l'inserimento di focus groups per coprire tutte le aree relative al: trasferimento di informazioni, alla preparazione del paziente e del caregiver, al supporto per l'autogestione e alla affermazione delle proprie preferenze. Si è scoperto che tale misura CTM ha la capacità di discriminare tra i pazienti che hanno avuto una visita al pronto soccorso o una re-ospedalizzazione per la loro condizione iniziale e coloro che non l'hanno avuta, ed è così in grado di far convergere i rapporti di esperienze negative riferite dai pazienti dopo la dimissione (Coleman, 2005). Questa misura è stata validata anche in altre popolazioni (Bakshi, 2012; Parry, 2008).

I pagatori pubblici e privati negli Stati Uniti e in altri paesi stanno riconoscendo l'esperienza del paziente come una componente fondamentale della valutazione della qualità delle cure erogate, e stanno tirando le somme dei risultati delle indagini sulle esperienze del paziente nelle strutture di providers sia pubblici che privati. Nel 2006, la Medicare Modernization Act ha legato i tassi di pagamento degli ospedali della rete Medicare per raccogliere e riportare pubblicamente i dati del *Consumer Assessment of Healthcare Providers and Systems* (CAHPS). La maggior parte degli ospedali statunitensi oramai raccoglie e pubblicamente diffonde tutte le domande standard dell'indagine CAHPS, e come gli ospedali statunitensi rispondono a questi standard di qualità (Report AHRQ, 2010).

Le opportunità di confronto tra PREMs

Nonostante le iniziative di raccolta dei PREMs che si svolgono a livello nazionale, non vi è ancora grande occasione di confrontare l'esperienza dei pazienti a livello internazionale. Diciannove paesi di area OCSE forniscono dati sull'esperienza del paziente con l'assistenza territoriale e i dati sono riportati in breve nel rapporto OCSE (*Health at a Glance*) in alcuni indicatori fra i più utilizzati.

Questi comprendono:

- il medico trascorre abbastanza tempo con i pazienti durante una visita;
- il medico fornisce spiegazioni di facile comprensione;
- il medico offre l'opportunità di porre domande e di esprimere preoccupazioni;
- il medico coinvolge i pazienti nei processi decisionali che interessano la loro cura e in generale il trattamento.

Tuttavia, l'ambito di questi indicatori è limitato al settore territoriale. L'OCSE non riferisce sull'esperienza del paziente negli ospedali, nei reparti di salute mentale, nei reparti di emergenza, nell'assistenza a lungo termine o per le cure palliative. Inoltre la relazione OCSE non tiene conto dell'esperienza del paziente per condizioni specifiche o per pazienti con polipatologia cronica.

Undici paesi partecipano alle indagini internazionali sulla politica sanitaria del Commonwealth Fund (CF).

Il CF conduce analisi internazionali sulla popolazione generale e anche valutazioni rivolte a gruppi di popolazione più specifici. Undici paesi partecipano al sondaggio del fondo: Australia, Canada, Francia, Germania, Paesi Bassi, Nuova Zelanda, Norvegia, Svezia, Svizzera, Regno Unito e Stati Uniti.

L'indagine sulla popolazione generale consiste in interviste telefoniche computerizzate, fatte su campioni casuali di adulti, utilizzando un questionario comune tradotto e adeguato per la formulazione specifica del paese e a seconda delle necessità

specifiche. Sulla base di quanto detto, una differenza fondamentale tra i sondaggi del Commonwealth Fund ed altri esistenti è che essi si estendono all'accesso e all'uso del Pronto Soccorso, ai tempi di attesa per vedere i medici, alle lacune nel coordinamento della cura e al costo come ostacolo all'assistenza sanitaria.

Le valutazioni PREMs attualmente svolte da paesi a livello nazionale, dall'OCSE e dal Commonwealth Fund a livello internazionale, sono una buona base per costruire una rete di valutazione PREMs a livello mondiale.

Una revisione strutturata delle indagini nazionali e trans-nazionali sull'esperienza del paziente per i paesi membri dell'Unione OCSE e non OCSE, intraprese a partire dal 1997 ha rilevato che, nonostante l'esistenza di più strumenti, pochi sono stati sperimentati per l'utilizzo in paesi diversi. Tra di essi troviamo lo World Health Organization's Health System Responsiveness study su 60 Paesi, condotto nel 2000-01 e il World Health Survey del 2002, che includeva anche domande sulla responsiveness dei sistemi (Garratt, 2008).

Sulla base dei più recenti lavori di letteratura, e i report dell'OCSE e del Commonwealth Fund, con la raccolta di molteplici indicatori in più settori e con lo studio di come i pazienti si muovano attraverso i percorsi di cura, si potrà avere l'opportunità di confronti tra paesi che potrebbero essere utili per i responsabili politici e contribuire a guidare miglioramenti delle prestazioni dei diversi sistemi sanitari. (Kyte, 2016a)

L'utilizzo di misure di esito del paziente (PROMs) nei sistemi sanitari

Un numero sempre maggiore di paesi sta mostrando interesse a utilizzare le misure di esito riferite dal paziente (PROMs) per saperne di più sugli outcomes di maggiore interesse per i pazienti. L'utilizzo di indicatori riportati dal paziente, insieme a altre metriche di qualità e outcomes più generici, fornisce un quadro più completo delle prestazioni erogate da parte del sistema sanitario, di cui possono beneficiare non solo i pazienti, ma anche i singoli medici, gli health-care providers e i policy makers. La raccolta di queste informazioni a livello internazionale è quindi fondamentale per la fornitura di servizi sanitari che ruotino sempre più attorno alle esigenze e alle domande di salute e di qualità della vita dei pazienti.

PROMs: la definizione generale di PROMs è la misura delle percezioni dei pazienti sul loro stato di salute, sugli esiti clinici, sulla mobilità e sulla qualità della vita.

A titolo di esempio possiamo citare le domande: *“quale era la mobilità del paziente prima di sostituire l'anca, ed è migliorata dopo l'intervento?”* oppure *“la condizione di un paziente limita la sua capacità di compiere attività faticose quali jogging, sci o ciclismo?”*

Tuttavia, le opportunità di confronto tra paese e paese sono ancora più limitate rispetto ai PREM. E questo fornisce ai responsabili politici solamente una comprensione superficiale dei risultati di performance dei servizi sanitari. I PROMs hanno quindi la potenzialità di migliorare tutte le pratiche cliniche.

Tipi di PROMs

Gli strumenti PROMs possono essere specifici per la malattia o generici. Quelli specifici per la malattia si concentrano sui sintomi e l'impatto sulla funzione di una condizione specifica (ad esempio se un paziente può andare a fare la spesa in modo indipendente dopo un intervento quale la chirurgia al ginocchio). Un esempio, invece, di un PROMs generico è l'EQ-5D, sviluppato dal gruppo EuroQol. I pazienti possono segnalare, ad esempio, dolore estremo o disagio, oppure ansietà. L'uso di strumenti specifici per le malattie e generici insieme può fornire una visione più completa degli outcomes su un paziente.

PROM generici relativi alla qualità della vita

EQ-5D

EQ-5D è una misura generica di esito della salute ed è applicabile a una vasta gamma di condizioni di salute e trattamenti, individuando 243 possibili "stati di salute".

Con questo questionario i pazienti descrivono il proprio stato di salute su cinque dimensioni:

1. mobilità,
2. auto-cura,
3. attività abituale,
4. dolore/disagio,
5. ansia/depressione

Per ognuna delle dimensioni è scelto un livello di gravità da 1 a 3: nessun problema, problema moderato o problema grave.

Il questionario EQ-5D genera quindi un indice EQ-5D che è un profilo sanitario che può essere trasformato in un indice globale di salute con un valore totale ponderato per la qualità della vita percepita, e rappresentando quindi una generale descrizione della

salute del paziente e, inoltre, confrontare come questo stato di salute è percepito dalla popolazione generale. Esistono, in realtà, diversi set di valori predefiniti per calcolare i punteggi e ogni set di valori rappresenta le preferenze della popolazione da cui è stata derivata. Quindi, i confronti dei risultati utilizzando indici di utilità calcolati con set di valori diversi (specie se sviluppati in diversi contesti culturali e/o geografici) possono essere difficili.

L'originale EQ-5D, che ha tre livelli di opzioni di risposta (EQ-5D-3L), è più comunemente usato ed è stato convalidato per i pazienti con osteoartrite. L'EQ-5D-5L è una versione estesa dell'EQ-5D che dispone di cinque opzioni di risposta per ogni dimensione. L'EQ-5D-5L possiede migliori proprietà psicometriche (come la migliore sensibilità) rispetto al EQ-5D-3L, ed è certamente previsto un 'aumento di utilizzo dell'EQ-5D-5L in futuri RCTs e registri di patologia.

Short Form 36 (SF-36)

L'indagine breve sulla salute o Short-form 36 comprende otto dimensioni della salute: il funzionamento fisico, il dolore fisico, le limitazioni al ruolo dovute ai problemi di salute fisica, le limitazioni al ruolo dovute a problemi personali o emotivi, il benessere emotivo, il funzionamento sociale, energia/stanchezza e percezione della salute in generale. L'SF-36 comprende anche un elemento per fornire un'indicazione della percezione del cambiamento nella salute. Gli otto punteggi scalari sono le somme ponderate delle domande nelle rispettive sezioni. Ogni scala viene poi trasformata direttamente in un punteggio da 0 a 100. Un punteggio pari a zero equivale alla massima disabilità e un punteggio di 100 equivale a nessuna disabilità. I punteggi delle otto scale sono riassunti in una componente fisica e una mentale su cui viene pesata la scala. Il SF-36 è il PROM generico più comunemente usato nelle sperimentazioni cliniche.

SF-12

SF-12 è una misura generica degli esiti di salute, composta da 12 elementi derivanti dal punteggio a 36 elementi, SF-36. SF-12 fornisce due punteggi riepilogativi: riepilogo

della componente fisica (Physical Component Summary o PCS) e riepilogo componente mentale (Mental Component Summary o MCS), che vanno da 0 a 100, in questo caso con i punteggi più alti ad essere migliori.

Per il follow-up di routine nei registri di interventi per artroprotesi SF-12 è considerato lo strumento gold standard rispetto all'SF-36. Gli strumenti SF richiedono la licenza, ma l'equivalente Veterans Rand 12-item survey (VR-12) e 36-item survey (VR-36) sono disponibili gratuitamente.

Crosswalk algorithms

Un recente studio ha valutato con successo un approccio probabilistico di mappatura sui punteggi di utilità EQ-5D basati sulle risposte SF-12 utilizzando reti bayesiane. Tale approccio di mappatura consente agli algoritmi crosswalk di convertire le risposte SF-12 ai punteggi dell'indice EQ-5D, e quindi, in buona sostanza, consentendo confronti tra i diversi strumenti.

Esempi di uso di registri clinici della qualità in alcuni paesi

La Svezia ha numerosi Registri Nazionali della Qualità (NQR) di cui 93 includono alcune forme di misure PROMs o PREMs (Nilsson, 2016). La metà di questi registri include alcuni tipi di misure generiche e più della metà include PROMs specifici per la malattia. Tra le misure generiche, quelle più comuni sono l'EQ-5D e l'SF-36 / RAND-36. L'uso principale dei PROMs in Svezia è stato in studi clinici o altre ricerche. L'uso dei PROMs nella pratica clinica e il miglioramento della qualità a livello locale sono in aumento. Esempi di NQR includenti misure PROMs sono il registro dell'artroprotesi dell'anca (EQ-5D), il registro della cataratta (Castquest-9SF) e il registro dello stroke (EQ-5D e altre misure).

È stato sviluppato un framework concettuale in cui inserire i registri australiani di qualità clinica al fine di stimolare il numero di raccolte dati capaci di catturare e riportare i dati di processi e risultati per specifiche condizioni cliniche o interventi. Lo sviluppo di una serie di registri nazionali ad alta priorità ha il potenziale per affrontare l'attuale lacuna nella misurazione della qualità dell'assistenza sanitaria e migliorare la qualità dell'assistenza ai pazienti. Il Framework, approvato dal Consiglio consultivo dei ministri della sanità australiana (AHMAC) nel marzo 2014, descrive le disposizioni nazionali per i registri di qualità clinica (ACSQHC, 2014). In Australia esistono diversi esempi di registri di esito nazionali - a volte in collaborazione con la Nuova Zelanda - come il Registro degli esiti per il cancro alla prostata (PCOR-ANZ). La raccolta dei dati include un questionario sulla qualità della vita (EPIC-26) a 12 e 24 mesi dopo il trattamento, attraverso telefonate di follow-up al paziente (Nag, 2016).

Nei Paesi Bassi esistono numerosi registri clinici che hanno iniziato a includere l'uso dei PROM. L'Istituto olandese per l'audit clinico (DICA) gestisce 23 registri nazionali, di cui nove registri comprendono i PROM (DICA, 2016). Esempi di registri, tra cui i PROM, sono il registro Dutch Surgical Spine Registry (Usa NDI o ODI) e quello della chirurgia

bariatrica (RAND36). DICA si sta muovendo per stimolare il miglioramento della qualità sulla base dei feedback che emergono dai dati.

L'uso dei PROMS in settori specifici

Chirurgia elettiva

La Svezia e i Paesi Bassi hanno sviluppato ciascuno i propri registri per l'artroplastica dell'anca, che raccolgono la totalità dei dati PROMs dei pazienti che si sottopongono a sostituzione artroprotesica dell'anca. Esiste anche un registro congiunto delle province della Nuova Zelanda che raccoglie dati PROMs da un campione random del 20% dei pazienti che effettuano un intervento chirurgico di sostituzione di una qualunque articolazione. Varie iniziative di raccolta dei dati PROMs esistono anche in diverse province del Canada. Ad esempio, l'Alberta ha il progetto pilota Edmonton Heart e Lung Transplant Clinic. Le informazioni sulla salute pre-operatoria vengono raccolte quando un paziente viene inserito nella lista dei pazienti in attesa di trapianto e le informazioni postoperatorie sono raccolte ogni volta che il paziente è ammesso in reparto (CIHI, 2015). La Svezia ha anche un network, Swespine, come base per la garanzia e il miglioramento della qualità. In particolare, Swespine misura la qualità della vita e la disabilità per tutte quelle persone con traumi alla colonna vertebrale, con valutazioni di follow-up a uno, due e cinque anni dopo l'intervento (Strömqvist, 2013).

Servizi di salute mentale

Nei Paesi Bassi, il monitoraggio degli outcome di routine (ROM) è usato per migliorare la qualità della cura della salute mentale. ROM è la misurazione sistematica degli esiti del trattamento nella pratica clinica di routine, e può essere utilizzata dai pazienti e dai medici per monitorare il progresso del trattamento. Può aiutare a determinare la diagnosi psichiatrica, i sintomi e il funzionamento psicosociale in ogni fase del trattamento.

I dati ROM, anonimizzati, possono essere utilizzati per condurre ricerche epidemiologiche e per scopi di benchmarking. La scala di valutazione della depressione

di Hamilton e la scala di valutazione della depressione di Montgomery-Åsberg sono esempi di scale di rating specifiche per le patologie mentali, che misurano la gravità del sintomo nella depressione maggiore e che sono utilizzate in questa esperienza olandese (van Noorden, 2013).

Degenza a lungo termine

InterRAI è una rete senza scopo di lucro di ricercatori in oltre 30 paesi. Mira a promuovere pratiche e politiche cliniche basate sull'evidenza attraverso la raccolta e l'interpretazione di dati relativi alle caratteristiche cliniche dei pazienti e agli esiti di salute in una varietà di contesti diversi. Gli strumenti di valutazione sono stati sviluppati per i diversi settings, comprese le strutture di assistenza a lungo termine. In misura minore, InterRAI è stato utilizzato in contesti di assistenza domiciliare. Gli indicatori InterRAI di qualità aggiustati per rischio e per tipo di provider comprendono: mobilità, comportamento, continenza della vescica, umore e dolore (Carpenter and Hirdes, 2013).

Gli strumenti InterRAI sono somministrati da valutatori che effettuano osservazioni basate su interviste con pazienti, oltre a consultare operatori sanitari e personale e riesaminare tutti i tracciati record e le cartelle cliniche disponibili. Ad esempio, per determinare se una donna anziana prepara i suoi pasti o riesce a svolgere i lavori di casa in un contesto di assistenza domiciliare, l'esaminatore consulterà la persona, il suo caregiver, i fornitori di servizi comunitari e qualsiasi documentazione disponibile. Inoltre, una serie di item sono specificatamente indirizzati alla persona che viene valutata, come ad esempio aspetti dell'umore e percezione della salute auto-riferita (Gray et al., 2009). Tra gli strumenti di indagine di InterRAI vi sono una serie di strumenti di qualità della vita soggettiva specifici per sito (QoL), progettati per offrire alle persone arruolate nei programmi di cure formali l'opportunità di esprimere le loro percezioni in ambiti quali: relazioni, ambiente, comfort, cibo e partecipazione alle attività quotidiane significative. Questi strumenti PROMs sono destinati a essere utilizzati durante le interviste con

pazienti che hanno la capacità cognitiva sufficiente per rispondere appropriatamente, e possono anche essere completati in tranquillità e autonomamente dalla persona e restituiti per posta (InterRAI, 2015).

Esistono anche questionari specifici per alcune malattie. Questi includono il questionario per la malattia di Parkinson (PDQ-39). Esso copre le dimensioni della mobilità, attività quotidiane di vita, benessere emotivo, stigma, supporto sociale, cognizioni, comunicazione e disagio fisico (Jenkinson, 1997). Per i pazienti affetti da demenza, lo strumento DEMQOL e DEMQOL-Proxy sono stati utilizzati rispettivamente per i pazienti e le persone che si prendono cura di loro. È stato trovato che DEMQOL e DEMQOL-Proxy forniscono un metodo efficace per valutare la salute e la qualità della vita e si raccomanda di utilizzarli insieme in caso di demenza lieve e moderata. Per i pazienti con grave demenza, solo il DEMQOL-Proxy è raccomandato (Smith, 2005).

Cure palliative

Il Palliative Care Outcome Scale (POS) è stato sviluppato per pazienti con tumore avanzato. Consiste di due misure quasi identiche, una completata dal personale, l'altra dai pazienti. Esso valuta una gamma di domini che includono dolore, ansia del paziente e dei familiari, sostegno, informazione, sensazione che la vita sia ancora utile, sentimento di autostima, tempo sprecato a causa degli appuntamenti relativi alla propria assistenza sanitaria (ad esempio, in attesa del trasporto fra strutture o casa/struttura o tempo perso per test ripetuti), se sono stati affrontati problemi pratici finanziari o personali derivanti dalla malattia e quali possono essere i principali problemi del paziente nei tre giorni precedenti (Hearn, 1999). Il POS è stato utilizzato in paesi tra cui Germania e Austria (Bausewein, 2005).

Informal care

I caregivers di pazienti con condizioni a lungo termine sono a rischio di alti livelli di stress, depressione e ansia, rendendo importante valutare la loro qualità di vita. L'indagine sulla salute in forma abbreviata (SF-36) è stata utilizzata per valutare la qualità della vita dei caregivers in diversi paesi con diversi gruppi di popolazione ed è stata tradotta e convalidata in molte lingue. Uno studio britannico su caregivers di pazienti con malattia dei motoneuroni, sclerosi multipla e morbo di Parkinson ha utilizzato l'SF-12 (derivato dall'SF-36), così come l'indice Carer Strain e un questionario di recente sviluppo sulle esperienze di assistenza sanitaria e sociale. Lo studio ha confermato che il benessere dei caregivers era compromesso nella maggioranza dei casi (Peters, 2013). In uno studio in Spagna, i ricercatori hanno tentato di utilizzare l'SF-36 in coloro che si prendono cura di pazienti con demenza. Lo studio ha rilevato che il 37,6% delle donne caregivers indicava come la propria salute fosse peggiorata o fosse molto peggiore rispetto all'anno precedente, rispetto al 26,2% delle donne nel gruppo di controllo. Nei maschi, le differenze non erano statisticamente significative (21,6% dei caregivers maschi, rispetto al 20,7% dei maschi nel gruppo di controllo). Tuttavia, gli autori hanno notato che una limitazione dello studio era che l'SF-36 non era stato convalidato rispetto ai caregivers di pazienti con demenza (Argimon et al., 2004). Una limitazione di entrambi gli studi è che si tratta di indagini trasversali.

Performance measurement statali per diverse condizioni patologiche

La Minnesota Community Measurement è un'iniziativa a livello statale negli Stati Uniti per la misurazione delle prestazioni sanitarie che si occupa di pubblicare i tassi di prestazioni sulle misure di qualità a livello statale e a livello dei gruppi di medici con regolarità. Il rapporto annuale sulla qualità sanitaria comprende ben 32 misure per condizioni diverse. Le PROMs sono incluse per diverse condizioni come la chirurgia spinale (variazione media dell'indice di disabilità Oswestry (ODI) a tre mesi post-intervento), sostituzione totale dell'articolazione del ginocchio (variazione media del

punteggio di Oxford Knee Score (OKS) a un anno dopo l'operazione) e depressione (remissione percentuale a sei mesi sulla base del Patient Health Questionnaire (PHQ-9)) (Snowden, 2015).

Uso dei PROMs

In un sondaggio e in una serie di interviste strutturate commissionate dal NHS inglese, esperti nazionali hanno confermato che l'uso di misure di esperienza segnalate dai pazienti (PREMs) è più diffusamente utilizzato rispetto all'uso delle PROMs. I PREMs sono stati usati per molti anni per misurare e migliorare la qualità delle cure. L'uso dei PROMs per scopi diversi dalla ricerca è relativamente nuovo.

Data la maggiore diffusione delle indagini PREMs, in diversi paesi (Australia, Canada, Paesi Bassi) le domande relative ai PROMs sono state aggiunte in coda ai questionari per le indagini PREMs, in precedenza o in concomitanza con l'attuazione dei programmi PROMs. Questo è il caso nello stato australiano del New South Wales, ad esempio, che ha aggiunto alcune domande per valutare i PROMs alle valutazioni PREMs per gli ospedali. Gli esperti nazionali hanno anche sottolineato che i PROMs e i PREMs servono obiettivi sia distinti che comuni. Le differenze nel modo in cui vengono somministrati PREM e PROM (ad es. tempistiche o frequenza) dovrebbero essere prese in considerazione nel valutare le opportunità e la fattibilità di somministrarle insieme. Infatti i PROMs sono utilizzati principalmente per la ricerca e per facilitare il processo decisionale condiviso tra medici e pazienti, per migliorare la pratica clinica e, a volte, fornire altri scopi per il miglioramento della qualità e per fare public reporting.

Chirurgia elettiva, salute mentale e cancro sono le malattie/condizioni più frequentemente considerate per le quali sono attualmente raccolti i PROMs. I PROMs sono stati talvolta utilizzati nell'assistenza a lungo termine, nelle cure palliative e nelle cure preventive.

Nella cura del cancro sono stati identificati due principali PROMs nel cancro alla prostata: il European Organization for Research and Treatment of Cancer, Quality of Life Questionnaire-C30 (EORTC-QLQ-C30); e l'Expanded Prostate Cancer Index Composite (EPIC-26).

La BPCO è la malattia cronica più comune per la quale vengono utilizzati PROMs, tra cui il test di valutazione della BPCO (CAT) e il questionario clinico per la BPCO (CCQ) sono i più frequentemente utilizzati. Infine, una varietà di PROMs specifici per la malattia vengono utilizzati nella cura della salute mentale, tra cui lo strumento Generalised Anxiety Disorder (GAD-7), la Manchester Short Assessment for Quality of Life (MANSA), la Depress Anxiety Stress Scale (DASS), il Patient Health Questionnaire (PHQ-9), la Symptom Checklist 90 (SCL90) e l'Hospital Anxiety and Depression Scale (HAD).

I PROMs generici sono stati talvolta utilizzati insieme a strumenti specifici per la malattia, per fornire un quadro più completo dei risultati dell'assistenza. In questi casi, lo strumento PROM generico più comunemente riportato dagli esperti nazionali è stato l'EuroQol (EQ-5D). Altri PROMs utilizzati comprendono le versioni dello Short Form (SF-36 e SF-12), l'Health Assessment Questionnaire (HAQ) e strumenti PROMIS generici. L'EQ-5D è sempre il più comunemente usato nella chirurgia ortopedica.

Alcuni studi recenti dimostrano che l'uso delle PROMs nelle patologie croniche è più difficile che per la chirurgia elettiva, con maggiori costi amministrativi e sfide metodologiche. La misurazione dei risultati di salute dei pazienti rispetto al coordinamento dell'assistenza e all'assistenza integrata è rara. (Kyte, 2017)

Nei paesi che stanno sperimentando l'utilizzo e la raccolta dei PROMs, essi sono più comunemente usati per misurare i risultati dei pazienti dopo un intervento chirurgico elettivo, in particolare le procedure di protesi articolare. Vi sono molte meno informazioni sui risultati dei pazienti con malattia cronica e pazienti trattati nei dipartimenti di salute mentale e assistenza a lungo termine. Ciò è dovuto principalmente al fatto che la natura cronica di queste condizioni presenta sfide nella progettazione e nella rilevazione dei dati delle indagini portate avanti con queste misure di esito. Inoltre, si sa molto poco sui risultati nelle aree dell'emergenza, dell'informal care, cure palliative e cure preventive. Per ciò che riguarda comunque le malattie croniche (e in particolare il

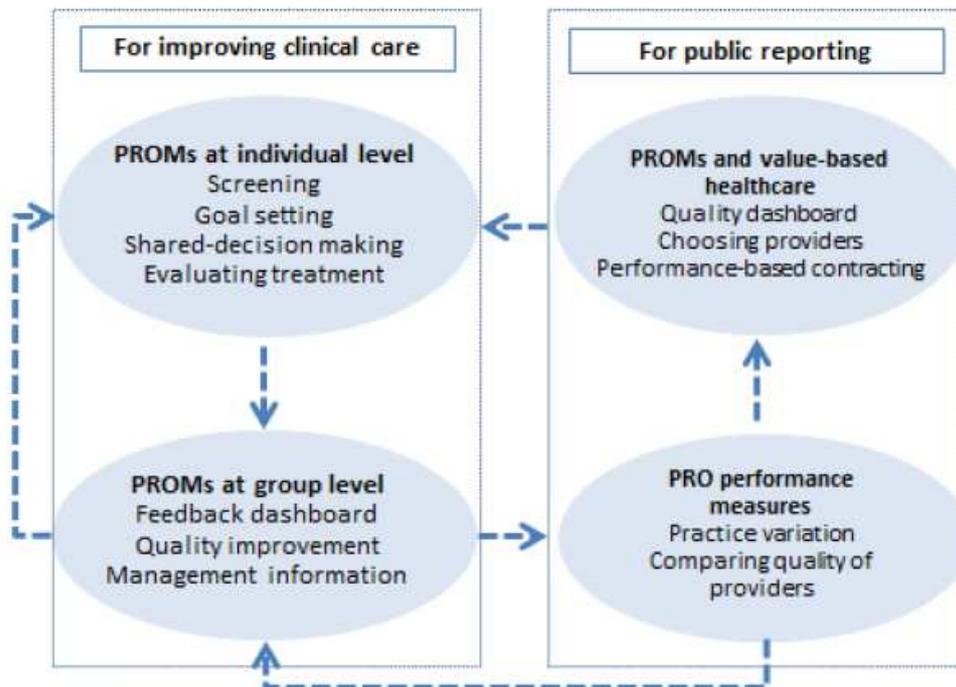
cancro), dato che questi pazienti eserciteranno un carico più significativo sul sistema sanitario nei prossimi anni, acquisire una migliore comprensione dei loro esiti diviene di fondamentale e sempre maggiore importanza.

Le procedure più frequentemente valutate coi PROMs sono la chirurgia del ginocchio (n = 9 paesi), la chirurgia dell'anca (n = 8 paesi) e la chirurgia della cataratta (n = 4 paesi).

I PROMs più comunemente utilizzati per la chirurgia dell'anca sono l'Hip Injury and Osteoarthritis Outcome Score – Physical Function Shortform (HOOS-PS) e l'Oxford Hip Score (OHS). I PROMs comunemente utilizzati per la chirurgia del ginocchio sono il Knee Injury and Osteoarthritis Outcome Score – Physical Function Shortform (KOOS-PS) e l'Oxford Knee Score (OKS). Nella chirurgia della cataratta tutti e quattro i paesi usano Catquest 9-SF. (Report OECD, 2017)

La Figura 1 presenta l'utilizzo dei PROMs nella pratica clinica, nel miglioramento della qualità e nella misurazione delle prestazioni. La figura mostra l'uso dei PROMs per uso interno da parte delle organizzazioni di servizi sanitari per migliorare la cura clinica; e per l'utilizzo esterno da parte delle parti interessate nella misurazione delle prestazioni e nel reporting pubblico.

Figura 1 Cornice concettuale per l'utilizzo dei PROMs



Source: Philip van der Wees, Radboud University Nijmegen.

Da questa figura evinciamo quindi come l'utilizzo dei PROMs si estrinsechi dentro le due cornici concettuali del miglioramento della cura rivolto ai pazienti e quella del reporting pubblico. Il miglioramento della cura passa sia a livello individuale che di gruppo. Il feedback del gruppo può essere estrinsecato in misure di performance che servono al reporting pubblico. Questo permette di creare una salute basata sul valore (fornito dai pazienti stessi attraverso i PROMs), che va a dare ricadute misurabili soprattutto per la salute del singolo e per favorire anche le sue scelte.

Per valutare l'utilizzo dei PROM e per promuovere la qualità, recenti report hanno effettuato revisioni di revisioni e hanno individuato inoltre singoli studi primari più recenti. In generale sono state individuate trenta revisioni, tra cui dodici revisioni quantitative e una revisione qualitativa. In sintesi, le revisioni mostrano che il feedback

dei dati PROMs nella pratica clinica può portare a miglioramenti nella qualità della cura del paziente. In particolare proprio nella cura oncologica esistono forti evidenze per l'utilizzo dei PROMs per individuare i sintomi, per migliorare la comunicazione tra i medici e i pazienti e infine per soddisfare i pazienti. Al contrario, ad esempio, nell'ambito della sanità mentale, non sono state trovate evidenze sufficienti per supportare l'uso routinario dei PROMs per migliorare la pratica clinica. (Report OECD, 2017)

I PROMs sono usati nella pratica clinica per aiutare a valutare i sintomi dei pazienti, la funzione e la qualità della vita e rispondere e migliorare tutti questi esiti con trattamenti adeguati.

I PROMs sono anche utilizzati per determinare l'efficacia dei trattamenti. Ad esempio, in uno studio, venivano presi in considerazione focus-groups di pazienti sopravvissuti a ictus, e in esso venivano identificati come risultati più rilevanti e significativi: l'essere sopravvissuti in casa, senza ictus ricorrenti e senza essere ricoverati per complicazioni. Lo studio ha quindi esaminato l'associazione tra trattamento con warfarin e esiti suddetti nei pazienti con fibrillazione atriale dopo ictus ischemico, rispetto ai pazienti senza anticoagulante orale dopo la dimissione. I pazienti trattati con warfarin in dimissione hanno avuto un rischio significativamente più basso di eventi avversi maggiori di natura cardiovascolare nei due anni successivi e soprattutto avevano un maggiore rischio di trascorrere più giorni di vita fuori dall'ospedale rispetto a quelli che non avevano ricevuto anticoagulanti orali alla dimissione (Xian et al., 2015). Lo studio è un esempio di come la ricerca focalizzata su outcome molto importanti per il paziente possa anche essere usata per sostenere le decisioni migliori da prendere sia da parte dei medici sia da un punto di vista scientifico e clinico.

I PROMs possono anche contribuire a migliorare la comunicazione tra pazienti e medici. In uno studio su pazienti affetti da tumore, quelli che hanno completato la raccolta dei dati dei PROMs discutevano più sintomi in modo significativamente maggiore durante le consultazioni con oncologi; in particolare dolore, stanchezza,

nausea e vomito. Tuttavia, il feedback dato dai PROMs non ha avuto alcun impatto sulla discussione più generali sul funzionamento nella vita di tutti i giorni e l'abilità a portare a termine le attività quotidiane (Takeuchi et al., 2011).

Un altro studio che ha indagato sulle cure reumatologiche in età pediatrica ha monitorato la qualità della vita relativa alla salute utilizzando PROMs elettronici. Questo studio ha trovato che l'uso dei PROMs aumentava la discussione di argomenti psicosociali e la soddisfazione dei reumatologi pediatri sul tipo di cura fornita durante la consultazione. Anche i genitori dei bambini hanno valutato positivamente lo strumento (Haverman et al., 2013). In uno studio su pazienti con schizofrenia, la misura della qualità della vita ha avuto un impatto positivo sulla soddisfazione del paziente, ma non ha alcun effetto sul miglioramento dei risultati clinici. Quest'ultimo dato potrebbe anche suggerire che i medici non abbiano utilizzato i dati in modo ottimale o che vi sia una effettiva complessità nell'utilizzare questo tipo di misure di esito sui pazienti psichiatrici (Boyer et al., 2013).

Raccolta e pubblicazione dei dati PROMs

A livello nazionale, la raccolta e la segnalazione dei dati riportati dal paziente possono informare i decisori politici in termini di allocazione delle risorse, individuare le variazioni e le disuguaglianze e aiutarli ad essere più sensibili alle esigenze dei pazienti. Inoltre, confrontare i risultati dei pazienti che utilizzano diversi servizi sanitari a livello nazionale può portare all'identificazione degli outliers che richiedono attenzione. Pubblicizzare pubblicamente le performance sulle metriche di questo tipo può fornire al pubblico informazioni importanti per aiutarli a fare le scelte migliori sui trattamenti e sugli ospedali. Questa trasparenza è fondamentale per un sistema sanitario veramente centrato sul paziente, anche se saranno necessarie campagne educative per generare la consapevolezza del pubblico sull'esistenza di queste informazioni, il loro scopo e, infine, la loro utilizzazione. In tempi recenti ci si chiede infatti, quanto l'attuale reporting pubblico dei dati sia in effetti usato dai pazienti. In generale, in letteratura, esiste molta poca evidenza a sostanzare come le scelte dei pazienti nella scelta del provider, siano effettivamente indirizzate dalla lettura attenta e dalla valutazione da essi data sui dati del reporting pubblico. La task force del National Health Service inglese, in particolare, ha individuato una serie di criticità, tra cui:

(i) i pazienti non sanno che i dati dei PROMs nazionali sono disponibili per questo scopo; (ii) i dati disponibili tramite il sito del Health & Social Care Information Centre non sono consegnati in un formato accessibile ai pazienti, quindi (iii) altri aspetti influenzano la scelta del paziente come: la raccomandazione del medico, la raccomandazione dei pari e soprattutto la località.

Questi risultati sembrano essere in accordo con la ricerca esistente. Una revisione sistematica recente, infatti, non è riuscita a trovare prove conclusive del fatto che la pubblicazione dei risultati delle prestazioni sanitarie fosse un fattore determinante del

comportamento dei pazienti in merito alla selezione del provider di cura. (Shekelle, 2008)

Sarà inoltre necessario compiere ulteriori ricerche per valutare le barriere al raggiungimento di questi obiettivi, in modo che la raccolta e la segnalazione di queste metriche si traducano in azioni che siano in grado di migliorare i risultati e indirizzino davvero le scelte consapevoli dei pazienti.

A livello organizzativo, vi è un chiaro potenziale dall'uso di PROMS. Ad esempio, gli ospedali del Bupa Systems, in Inghilterra, hanno iniziato a raccogliere i dati di PROMS nel 1998, originariamente motivati, fondamentalmente, dal desiderio di identificare le "mele marce" da un punto di vista clinico. Tuttavia, su questa scorta, i ricercatori che hanno lavorato nella rete Bupa hanno anche riconosciuto che i PROMS hanno potuto condurre a un miglioramento costante della qualità dei servizi erogati, hanno potuto cambiare la pratica clinica e hanno fornito risposte ai medici e ai pazienti. Inoltre, pubblicando i risultati dei livelli dei PROMS dei loro ospedali sui loro siti web, gli ospedali Bupa hanno potuto promuovere i benefici sulla qualità della vita ottenuti dalle prestazioni e dagli interventi che hanno fornito nel corso degli anni (Devlin et al., 2010). Bupa si è concentrata su procedure a più alto volume e ad alto rischio per assicurare che i risultati fossero statisticamente rilevanti. I dati dei PROMS sono stati utilizzati come un'indicazione di dove altri dati di processo e risultati potrebbero richiedere un'analisi più approfondita. In un esempio di come Bupa ha utilizzato i dati PROMS per promuovere la migliore pratica possibile, è stato identificato un ospedale della rete che aveva un aumento significativamente superiore, in termini di salute dei pazienti dimessi dopo la sostituzione dell'anca rispetto alla media dei loro ospedali. L'analisi delle caratteristiche di questo ospedale ha rivelato come il reparto di fisioterapia avesse avviato un intenso lavoro preoperatorio dei pazienti che intendevano sottoporsi a chirurgia sostitutiva dell'anca, ciò significava che i pazienti erano meglio preparati sia per la loro procedura chirurgica che per il successivo periodo di recupero. Tale pratica è

stata poi adottata da tutti gli ospedali della rete, migliorando gli esiti per tutti i pazienti (Devlin et al., 2010).

Mentre l'uso di PROMs per migliorare la pratica clinica è ancora relativamente nuovo, tuttavia esistono molti esempi di utilizzo a questo proposito. L'International Consortium for Health Outcomes Measurement (ICHOM) è consapevole di circa 185 organizzazioni che implementano uno standard set ICHOM. In particolare 33 organizzazioni che implementano almeno più di un set e 14 registri di malattia che ne misurano almeno uno. ICHOM supporta direttamente 40 organizzazioni attraverso l'implementazione di standard codificati al loro interno. Ad esempio, la Aneurin Bevan Health Board del Galles ha recentemente implementato il set di indicatori standard per la malattia di Parkinson e si prepara ad implementare anche i set per il mal di schiena e per la cataratta (ICHOM, 2015).

Gli esempi della letteratura scientifica pubblicata suggeriscono quindi che la raccolta e la segnalazione dei dati PROMs siano in grado di migliorare la pratica clinica e il processo di cura. Idealmente, i providers, sia a livello organizzativo che individuale, dovrebbero utilizzare le informazioni per individuare dove hanno bisogno di migliorare, perché in qualche caso potrebbero anche non far così bene come fanno altri loro concorrenti. Inoltre, tramite i PROMs, potrebbero individuare quali strategie possono adottare per ottenere miglioramenti sostanziali. Se da un lato, quindi, la letteratura ha individuato molti punti a sostegno dell'implementazione dei PROMs per il miglioramento delle policy sanitarie, le prove sull'impatto dei PROMs sugli esiti dei pazienti rimangono limitate.

Una revisione dell'uso dei PROM nella pratica clinica del cancro suggerisce inoltre che è necessaria fare ulteriori ricerche per tradurre le conoscenze acquisite dai PROMs in azioni efficaci che possano realmente modificare il comportamento per ottenere un effetto sugli esiti di salute (Howell et al., 2015).

La raccolta dei PROMs

La raccolta dei PROMs a livello di sistema è ancora in uno stadio molto embrionale. L'NHS in Inghilterra è stato il primo sistema sanitario del mondo a introdurre la raccolta di dati PROMs a livello di sistema. Dal 2009 il Department of Health ha richiesto la misurazione di tutti i pazienti in Inghilterra prima e dopo il ricovero in caso di quattro possibili procedure elettive cui il paziente poteva essere sottoposto. I pazienti hanno completato sia le indagini generiche che quelle specifiche per la chirurgia di artroprotesi di ginocchio, dell'anca e la chirurgia per vene varicose. Per la riparazione dell'ernia al disco, i pazienti completavano solo un'indagine generica (tabella 1.1).

Tabella 1: Programma PROMs nel sistema NHS, UK

Treatment	Condition-specific PROM	Generic PROM
Knee replacement	Oxford Knee Score	EQ-5D (including EQ VAS)
Hip replacement	Oxford Hip Score	EQ-5D (including EQ VAS)
Varicose vein removal	Aberdeen Varicose Vein Questionnaire	EQ-5D (including EQ VAS)
Hernia repair	No instrument	EQ-5D (including EQ VAS)

Source: Health & Social Care Information Centre (2015). Note: EQ VAS = EQ Visual Analogue Scale.

L'obbligo alla raccolta dei PROMs per le quattro procedure suddette fa parte del contratto NHS per i servizi pubblici che curano le patologie acute, e ha lo scopo di migliorare la qualità clinica e gli esiti per i pazienti. Questo sistema è parte integrante del NHS Outcomes Framework, progettato per fornire la accountability degli esiti che l'NHS fornisce e per monitorare le prestazioni dei servizi sanitari. Le informazioni sono pubblicamente disponibili ai cittadini sul sito My NHS, consentendo i confronti a livello ospedaliero.

Tabella 2. Partecipazione al programma nazionale PROMs e tassi di linkage allo stesso, da aprile 2013 a marzo 2014. (Kyte, 2017)

Procedure	Total procedures conducted	Total episodes successfully linked to both a pre- and postoperative questionnaire (%)
Hip replacement	76,576	44,536 (58.2%)
Knee replacement	79,769	45,606 (57.2%)
Groin hernia	73,229	23,481 (32.1%)
Varicose vein	26,201	5050 (19.3%)

Il costo totale del programma è stato stimato a 825.000 sterline l'anno. (Report NHS, 2016) Tuttavia, non è chiaro se questa spesa abbia prodotto cambiamenti quantificabili nel comportamento dei fornitori e nei benefici associati per i pazienti. (Varanguam 2014) Ci sono anche preoccupazioni riguardo a tassi preoccupanti di dati mancanti. Ad esempio, nel 2013/2014, la proporzione di episodi linkati con successo sia a un questionario pre- che post-operatorio (vale a dire dove è possibile calcolare il guadagno di salute del paziente a livello aggregato) aveva un range oscillante dal 58,2% al solo 19,3%

Mentre l'Inghilterra è la più avanzata nell'implementazione di un programma di PROMs a livello di sistema, un certo numero di altri paesi mostra interesse a raccogliere dati sulle PROMs in settori quali chirurgia elettiva, assistenza sanitaria mentale, assistenza a lungo termine, cure palliative e informal care.

Impatto dei PROMS

L'impatto del feedback fornito dai PROMs sui risultati delle cure è limitato, tuttavia, sottolinea l'importanza di ulteriori attività e ricerca in questo settore. L'esperienza e l'evidenza che emerge dalla esperienza del NHS segue questa considerazione. Infatti nello studio del 2013 che ha analizzato i dati (Varagunam, 2014) si è definito come l'iniziativa di raccogliere questo tipo di dati abbia portato, a livello dell'NHS, cambiamenti molto poco sostanziali, sia nelle differenze inter-providers, sia nel cercare di chiudere il gap fra i cosiddetti "outliers" e la media dei providers nel corso del tempo.

Anche dagli studi che emergono dalla letteratura scientifica le informazioni date dai PROMs sembrano, quindi, da sole insufficienti per individuare le opportunità e le criticità che possano far ottenere un effettivo miglioramento della qualità. In generale studi rigorosi sugli effetti dell'utilizzo di PROMs come strumento di misurazione delle prestazioni sono quindi ancora relativamente scarsi, e riflettono il numero limitato di programmi finora attuati e la base dei dati ancora poco sistematizzati sulla base di questo tipo di indicatori di performance riportati dai pazienti. Una revisione sistematica da parte di Boyce e Browne (Boyce, 2013) aveva identificato un solo studio di feedback sulle prestazioni a livello di gruppo e non ha trovato alcun effetto sulle prestazioni. Chen e colleghi (Chen, 2013) non hanno trovato studi che abbiano valutato l'utilizzo di PROMs in ambito oncologico per migliorare la qualità, la trasparenza, la responsabilità, il reporting pubblico o le prestazioni del sistema.

Un altro studio in chirurgia ortopedica di Boyce e Browne (Boyce, 2015) ha mostrato che i risultati per i pazienti operati da chirurghi che avevano ricevuto i dati di feedback PROMs sulla base di attività di benchmark fra pari non erano statisticamente diversi dagli esiti dei pazienti gestiti da chirurghi che non ricevevano feedback. Uno studio recentemente pubblicato da Varagunam (Varanguam, 2014) ha suggerito che le prestazioni ospedaliere non sono state alterate dall'introduzione delle misure di esito PROMs routinarie nell'ambito della chirurgia, in Inghilterra. Tuttavia gli autori, non

trovano riscontro di un impatto significativo sui processi di cura, hanno concluso che la modalità di comunicazione dei risultati, la necessità di un feedback tempestivo e l'inclusione di azioni suggerite per migliorare i PROMs potrebbero essere necessarie per ottenere un vero cambiamento del paradigma. Partridge (Partridge, 2016) ha condotto uno studio per il miglioramento di qualità, senza gruppo di controllo, nella sostituzione totale del ginocchio con artroprotesi e utilizzando dati PROMs. Le differenze statisticamente significative negli esiti e nel recupero - misurato con l'Oxford Knee Score – dopo l'intervento chirurgico a seconda delle diverse marche di impianto utilizzate era un motivo per cambiare e/o utilizzare le protesi più performanti e con recuperi migliori e più veloci.

Studi qualitativi mostrano che è complesso utilizzare i PROMs come strumento di miglioramento della qualità ed è talvolta necessario avere un feedback personalizzato per supportare l'interpretazione dei PROMs in modo da migliorare la qualità dell'assistenza erogata. Ciò diventa ancora più importante con l'utilizzo di dati aggregati per comprendere la variazione nei risultati tra i medici e le organizzazioni (Boyce, 2014; Howell, 2015). Questi risultati confermano gli studi sull'utilizzo dei dati dell'esperienza del paziente anche nella misurazione delle prestazioni (Friedberg, 2011). Alcuni autori hanno teorizzato come la potenziale carenza di impatto dei PROMs possa dipendere dalla fornitura inadeguata di un feedback ai providers, ai clinici, ai pazienti, non soddisfacendo, in buona sostanza, le esigenze degli utilizzatori finali dei PROMs (Varanguam, 2014)

In buona sostanza, quindi, tutti questi studi recenti (costruiti dopo l'implementazione in molti stati, in particolare UK e USA) suggeriscono la necessità di integrare la misurazione delle prestazioni in un programma formale e più generale di miglioramento della qualità. L'utilizzo, isolato, dei soli PROMs, è sostanzialmente insufficiente (Squitieri, 2017)

La Taskforce che si occupa di valutare il dato per NHS ha trovato che un'altra ragione per la quale non è facile ottenere un cambiamento vero con i PROMs è che i clinici

vanno persuasi sull'evidenza che è sottesa all'uso dei PROMs per dirigere un vero cambiamento dei servizi sanitari.

Nel 2014, l'NHS England ha implementato uno schema di "pay-for-performance" per la chirurgia della protesi d'anca e ginocchio, finalizzato a ridurre le variazioni fra providers. (Monitor & NHS England, 2014) Il pagamento della tariffa è subordinato al fatto che i providers dimostrino un guadagno medio di salute che non diminuisce in modo significativo sotto il tasso di 'allarme', definito come superiore a tre deviazioni standard al di sotto della media nazionale. I providers che non hanno rispettato questi criteri hanno ricevuto pagamenti per le procedure del 10% (£ 550) al di sotto della tariffa. Sebbene sia lodevole l'impegno a collegare la retribuzione del fornitore ai risultati centrati sul paziente, sembra chiaro che nella sua forma attuale, l'iniziativa tariffaria sulle migliori pratiche possa influenzare il comportamento di una piccola minoranza di trust. Utilizzando l'attuale soglia di tre deviazioni standard, ciò equivale a appena allo 0,1% dei fornitori.

Ulteriori conoscenze su come i PROMs vengono valutati, raccolti e implementati dai sistemi sanitari dovrebbero emergere da due revisioni della letteratura attualmente in corso. Una realistic review di Greenhalgh e colleghi che mira a capire con quali mezzi e in quali circostanze il feedback dei dati PROMs porta a effettivi miglioramenti del servizio (Greenhalgh, 2015).

D'altro canto, Concalves e colleghi stanno attualmente conducendo una revisione Cochrane con l'obiettivo di valutare l'impatto dell'utilizzo routinario dei PROMs nella pratica clinica, sul processo di cura; sulle esperienze di cura dei pazienti e dei professionisti; e sugli esiti di salute (Concalves, 2016).

Caso di studio: indicatori del cancro alla prostata

Il cancro della prostata è la diagnosi più comune di tumore maligno e la seconda causa di morte negli uomini. (Siegel, 2016) Gli ultimi 20 anni hanno visto un calo di oltre il 50% dei tassi di mortalità per cancro alla prostata in virtù di una migliorata diagnosi e trattamento precoce (Etzioni, 2012). Anche se il controllo sulla progressione del cancro e sui tassi di sopravvivenza dopo la prostatectomia sono eccellenti (Eggerer, 2011, ACS, 2016), i pazienti possono sviluppare effetti collaterali legati al trattamento.

Per il cancro alla prostata localizzato, la prostatectomia radicale (PR) è una delle molte scelte possibili disponibili. La decisione di adottare questo trattamento è basata sul tipo di cancro e la clinica, ma anche sull'età, le comorbidità e le preferenze del paziente (Cooperberg 2010). I dati di letteratura confermano come la PR sia il trattamento di prima linea più utilizzato per il cancro della prostata, con quasi il 50% degli uomini che si sottopongono inizialmente a questa tipologia di intervento. (Cooperberg 2010a).

Le variazioni nella cura del cancro alla prostata sono ben documentate. (Ghani, 2015) Il trattamento varia a seconda che siano coinvolte popolazioni vulnerabili, regioni geografiche diverse, ospedali e/o singoli clinici. (Hanan, 1995) Le metriche di qualità vanno ad indirizzarsi specialmente verso l'analisi di eventuali differenze di trattamento o su esiti diversi e includono misure di struttura, di processo e di esito per l'assistenza che viene erogata. (Donabedian, 1988) Negli ultimi dieci anni sono state proposte molte misure sulla qualità del cancro alla prostata, tra cui un set iniziale e completo di misure per il cancro alla prostata in fase iniziale (screening alla prostata) sviluppato da RAND. (Blumenthal, 2015)

Il numero di misure di qualità proposte, approvate e/o indirizzate agli operatori sanitari è aumentato rapidamente nell'ultimo decennio. Per la cura del cancro alla prostata, diverse misure di qualità sono utilizzate o approvate da varie agenzie di regolamentazione, sia private che pubbliche in tutto il mondo, sia per valutare la qualità sia per rimborso dei servizi. (Panzer 2013, Chien 2013) La pleora di possibili alternative

sugli indicatori di qualità è potenzialmente in grado di causare difficoltà nel conoscere su quali misure è preferibile mettere a fuoco i propri obiettivi. Alcune esperienze di audit, inoltre, ci sottolineano come molti operatori affermino che le molte misure di qualità (diverse fra loro) causino un onere notevole di lavoro, talvolta inutile, che a sua volta può causare/aumentare il burnout clinico. (Nowels 2016, Casalino 2017) Ciò suggerisce la necessità di identificare una serie di misure di qualità di base da mettere meglio a fuoco dato che, stante anche quanto enunciato nella parte iniziale di questa tesi, le misure di qualità costituiscono ad oggi una componente fondamentale e una parte imprescindibile del processo di miglioramento della efficienza. Un'adeguata misurazione della qualità della cura può fornire dati per prevenire l'uso eccessivo, l'abuso e il sottoutilizzo delle risorse sanitarie. Ma questo passaggio non può prescindere dal basare il proprio sistema anche sul valore dell'assistenza erogata, con un focus particolare sulle metriche basate/riferite dai pazienti. E che maggiormente, oltre che sul sistema, impattano sulla salute e sulla qualità della vita dell'individuo. E tutto questo passa attraverso la misurazione della qualità, anche poiché il sistema sanitario non può migliorare ciò che non può essere misurato o valutato. (IOM, 2011)

sperimentale della presente tesi sono stati identificati gli indicatori più importanti di qualità per il cancro alla prostata (proposti dalla comunità scientifica) e determinarne il loro utilizzo. Questo progetto ha, in particolare, tre obiettivi specifici:

- 1) sviluppare un elenco completo di tutte le misure di qualità disponibili per valutare la cura del cancro alla prostata. (le misure saranno classificate secondo la classica metrica Donabediana di “struttura”, “processo” e “esito”);
- 2) capire come sono state utilizzate queste misure, ad esempio per il miglioramento effettivo della qualità rispetto alla ricerca.
- 3) Identificare le misure di qualità più utilizzate in letteratura per la valutazione del tumore alla prostata;

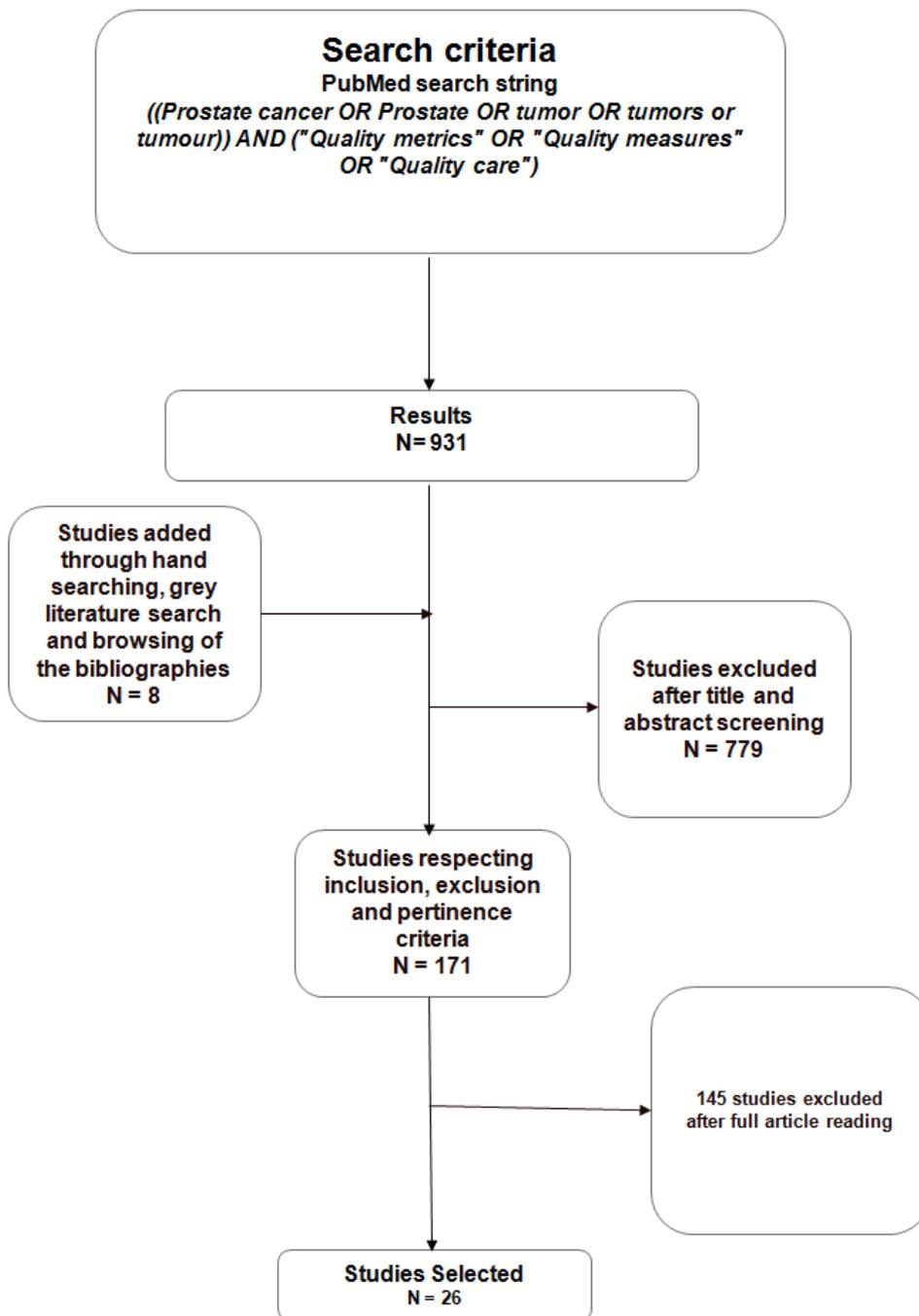
Metodi

Abbiamo consultato il database elettronico Medline (PubMed) utilizzando la stringa: *((Prostate cancer OR Prostate OR tumor OR tumors OR tumour)) AND ("Quality metrics" OR "Quality measures" OR "Quality care")*. La stringa di ricerca è stata creata combinando i termini relativi al cancro alla prostata e agli indicatori di qualità, utilizzando free-terms o MESH-terms di PubMed e combinandoli con gli operatori booleani più appropriati per essere il più sensibili e specifici possibile.

Tutti gli articoli pubblicati in inglese sono stati inclusi; non è stata adottata alcuna selezione basata sul tipo di pubblicazione. Non vi erano restrizioni temporali di data, e l'ultima ricerca è stata effettuata il 31 luglio 2017. Qualsiasi tipo di articolo primario è stato incluso nel processo di selezione e nella valutazione finale. Non sono stati inclusi articoli primari minori (ad esempio, letters to the editor, commentaries, ecc.) oppure articoli secondari/di revisione. Le bibliografie degli articoli di revisione e degli studi primari selezionati sono stati scansionati per trovare articoli aggiuntivi. Inoltre è stata effettuata una ricerca della letteratura grigia e ricerche manuali, comprensive di snowball, delle maggiori riviste, siti web di organizzazioni professionali e istituti che formulano linee guida. (Greenhalg, 2005)

La figura 1 mostra il diagramma di flusso per il processo di selezione degli articoli.

Figura 1: Flow chart che mostra il processo di selezione degli articoli



Le misure o i concetti proposti sono stati classificati principalmente sulla base di quante volte sono stati citati e/o valutati negli studi trovati e secondariamente sulla base delle dimensioni del campione in cui sono stati implementati/testati. Gli studi sono stati classificati come "Quality improvement" se è stato specificato nel loro scopo che erano sostanzialmente diretti verso il miglioramento della qualità delle cure. Al contrario, sono stati definiti come "Studi di ricerca" quelli che non avevano come obiettivo primario il miglioramento della qualità dell'assistenza diretta. Infine, abbiamo deciso di creare una tabella "top ten" generale delle misure/concetti più citati e utilizzati, e anche una tabella "top five", raccogliendo le cinque misure che forniscono le informazioni più dettagliate sulla fornitura di assistenza sanitaria di qualità nelle tre principali aree (struttura, processo, esito - secondo il modello di Donabedian) per il tumore della prostata.

Risultati

La ricerca iniziale ha prodotto 931 articoli. La prima selezione è stata effettuata attraverso lo screening dei titoli e degli abstract. 171 articoli sono risultati idonei per la lettura in-esteso. Dopo questa seconda fase, 26 articoli (Bekelman, 2007; Brundage, 2013; D'Avolio, 2008; Hernandez-Boussard, 2015; Herrel, 2017, Holmes 2017; Kamal 2013; Lee 2012; Makarov, 2012; Miller, 2007; O'Connor, 2014; Ranpura, 2015; Samuel, 2014; Schroeck, 2015; Schroeck, 2014A; Schroeck, 2014B; Schroeck, 2014C; Shelton, 2014; Skolarus, 2013; Sohn, 2016; Toole, 2012, Webster, 2012; Wei, 2002; Worwag, 1998; Wright, 2015) sono stati selezionati come articoli idonei per l'estrazione dei dati. La ricerca manuale e la ricerca nella letteratura grigia hanno aggiunto 8 articoli pertinenti.

La Tabella 1 riassume alcune delle caratteristiche descrittive delle misure di qualità della cura del cancro alla prostata valutate. Abbiamo identificato 71 misure di qualità proposte da questi articoli e dai siti web delle istituzioni/iniziative proponenti. Le misure

di processo rappresentavano l'84,5% di tutte le misure proposte. Sono state proposte solo 6 misure di esito e 5 misure strutturali, ciascuna delle quali contribuisce meno del 10% delle misure totali valutate. Abbiamo trovato 7 misure di qualità, dalla lista, che sono state approvate dal National Quality Forum (NQF). La RAND è stata l'istituzione leader nella proposta di misure di qualità del cancro alla prostata, poiché aveva proposto da sola ben 23 misure di qualità, pari a circa il 32,5% di tutte le misure valutate. Altre istituzioni/iniziativa proponenti erano la Quality Ontology Initiative (QOPI) della American Society of Clinical Oncology (ASCO), il registro AUA Quality (AQUA) della American Urological Association (AUA), il College of American Pathologists (CAP), il Physician Consortium for Performance Improvement (PCPI), l'American Medical Association (AMA) e il Physician Quality Reporting Initiative (PQRI).

Tabella 1. Caratteristiche descrittive delle misure di qualità per il carcinoma prostatico

Misura	Proposta da	Approvata da	Numero di referenza
Struttura			
Proporzione consistente di pazienti trattati da provider ad alti volumi (terzile più alto) – riferito al chirurgo e al radioterapista	RAND		4
Disponibilità di strutture di oncologia e radiologia e la presenza di una consulenza psicologica per i pazienti	RAND		4
Presenza nella struttura di una certificazione del board degli urologi e dei radiologi oncologi	RAND		4
Informazioni sugli esiti per i pazienti trattati dal provider	RAND		4
Prove di una aderenza istituzionale al protocollo di pratica del College dei Patologi Americani per la gestione dei campioni istologici di patologia	RAND		4
Processo			
DRE (Esplorazione digitale rettale)	RAND		4
Almeno 10 campioni tissutali prelevati al momento della biopsia prostatica			20
Documentazione dei livelli del PSA (Prostate Specific Antigen)	RAND		4
Documentazione clinica dello stadio T	PQRI		21
Valutare lo stadio della malattia prima della stadiazione clinica trattamento o pre-trattamento attraverso DRE, PSA totale, Gleason Grade	RAND		4

Valutazione della storia familiare di carcinoma prostatico	RAND		4
Valutazione delle comorbidità	RAND		4
Scansione ossea eseguita in pazienti ad alto rischio	NCCN		22
Scansione TC eseguita in pazienti ad alto rischio	NCCN		22
Evitare misure eccessive - Scansione ossea per la stadiazione di pazienti a basso rischio	AMA	NQF	7
Evitare misure eccessive - TAC per la stadiazione di pazienti a basso rischio	PCP		7
Una chiara descrizione del rischio di complicazioni del trattamento	RAND		4
Punteggio Gleason riferito al paziente	RAND		4
Valutazione del paziente del funzionamento urinario, sessuale e intestinale prima del trattamento	RAND		4
Percentuale di pazienti ad alto rischio sottoposti a linfoadenectomia	AQUA		45
Misurazione della quantità di perdita di sangue durante una prostatectomia radicale	RAND		4
Uso appropriato di antibiotici e interruzione	AQUA		45
Appropriati livelli di tromboprolifassi	AQUA		45
Uso di terapia ormonale adiuvante prima della brachiterapia	AQUA		45
DEXA Scan per uomini che iniziano ≥ 12 mesi di terapia ormonale	AQUA		45
Documentare il colesterolo e / o i livelli di glucosio prima di iniziare la terapia ormonale ≥ 12 mesi e di nuovo 12	AQUA		45

mesi dopo l'inizio			
Percentuale di utilizzo della sorveglianza attiva per i pazienti a basso rischio	AQUA		45
Documentazione ≥ 2 PSAx entro un anno dopo il trattamento o dall'inizio della sorveglianza attiva	AQUA		45
Documentazione del numero di cicli di chemioterapia	QOPI		45
Comunicazione riassuntiva del trattamento	AMA	NQF	7
Efficacia dei narcotici e antidolorifici valutata in visita dopo prescrizione	QOPI		46
Intenzioni sulla chemioterapia documentate (intento curativo vs palliativo documentato prima o entro due settimane dalla somministrazione)	QOPI		46
Documentazione della discussione degli intenti della chemioterapia con paziente	QOPI		46
Uso di fattori stimolanti delle cellule bianche quando il rischio di neutropenia febbrile per i pazienti è inferiore al 20%	QOPI		46
Consenso paziente firmato, per la chemioterapia	QOPI		46
Documentazione sul consenso del paziente nella nota del medico	QOPI		46
Riepilogo del trattamento chemioterapico completato entro 3 mesi dalla fine della chemioterapia	QOPI		46
Trattamento chemioterapico fornito al paziente entro 3 mesi dalla fine della chemioterapia	QOPI		46
Riassunto del trattamento chemioterapico fornito o comunicato al (i) professionista (i) entro 3 mesi dalla fine della chemioterapia	QOPI		46

Stato di tabagismo documentato prima della seconda visita di controllo (stato di tabagismo / uso del tabacco documentato per l'anno precedente dai questionario QOPI)	QOPI		46
Visita per smettere di fumare raccomandata ai fumatori di sigarette entro la seconda visita di controllo	QOPI		46
Azione intrapresa per affrontare i problemi con il benessere emotivo dalla seconda visita	QOPI		46
Tasso di complicanze chirurgiche acute	RAND		4
Uso della tomografia computerizzata per pianificare il trattamento radioterapico	RAND		4
Immobilizzazione del il paziente durante il trattamento, proteggendo la mucosa rettale	RAND		4
Segnalazione di patologie/complicanze correlate alla prostatectomia radicale	CAP	NQF	7
Terapia di deprivazione androgenica (ADT)/terapia ormonale adiuvante per pazienti con tumore prostatico ad alto rischio	AMA	NQF	7
Chemioterapia a base di docetaxel per carcinoma prostatico resistente alla castrazione ormonale e/o metastatico			35
Radioterapia ricevuta (RT) entro 30 giorni dalla morte			37
Seguire il paziente almeno due volte durante il primo anno di trattamento	RAND	NQF	4
Chirurgia eseguita in anestesia epidurale tramite prostatectomia pubica retroperitoneale			31
Ha ricevuto la chemioterapia negli ultimi 14 giorni di vita	RAND		4

CRT 3-D o IMRT per pazienti con carcinoma della prostata trattati con EBRT			36
Uso di fotoni ad alta energia (> 10 mV)			44
Dosi sull'asse centrale di almeno 75 Gy durante la radioterapia			35
Prescrizione di farmaci antidolorifici per malati di cancro in fase avanzata			36
Comunicare con il medico di base del paziente per garantire continuità assistenziale	RAND		4
Piano di cura per il dolore moderato/severo documentato	QOPI	NQF	7
Dolore valutato dalla seconda visita	QOPI	NQF	7
Stitichezza valutata al momento della prescrizione di narcotici o nella visita successiva	QOPI		25
benessere emotivo del paziente valutato dalla seconda visita	QOPI		25
Monitoraggio del PSA dopo il trattamento			43
Monitoraggio con DRE durante la sorveglianza attiva			43
Ripetizione della biopsia ogni 18 mesi nella sorveglianza attiva			43
Uso di sorveglianza attiva, <i>watchful waiting</i> per il cancro alla prostata a basso rischio	AQUA		21
Esito			
Ricovero in ospedale o trattamento medico o chirurgico per una serie di gravi complicanze	RAND		4

Evitare ricoveri ospedalieri multipli negli ultimi 30 giorni di vita			27
Ammissione all'Hospice negli ultimi giorni di vita			28
Margini chirurgici ben conservati	AQUA		23
Valutazione del funzionamento urinario, sessuale e intestinale del paziente dopo il trattamento	RAND		4
Soddisfazione dei pazienti per la scelta del trattamento, per la continenza e la potenza sessuale	RAND		4

La Tabella 2A illustra l'uso delle misure di struttura in letteratura. Delle 5 misure identificate, solo 3 sono state utilizzate in studi scientifici. La misura strutturale più frequentemente utilizzata era la *proporzione di pazienti trattati da un provider ad alti volumi*, ed è stata pubblicata in 6 studi. Un'altra misura strutturale, *la presenza nella struttura di una certificazione del board degli urologi e dei radiologi oncologi*, veniva utilizzata in 2 studi. La terza misura di struttura, *la disponibilità di strutture di oncologia e radiologia e la presenza di una consulenza psicologica*, appariva solamente in uno studio.

La tabella 2B riporta 38 misure (circa il 63%) di tutte le misure di processo utilizzate in letteratura. Il 20% di queste misure di processo, tuttavia, erano utilizzate solo in un solo studio.

La Tabella 2C illustra gli usi delle misure di esito, quelle di nostro maggiore interesse. Tutte e sei le misure di esito proposte sono state utilizzate negli studi da noi analizzati. Tuttavia, solo 2 su 6 misure sono state utilizzate in più di uno studio. Esse, in particolare, erano due PROMS riguardanti:

1. la valutazione del funzionamento urinario, sessuale e intestinale del paziente dopo il trattamento;
2. la soddisfazione del paziente per la scelta del trattamento, la continenza e la potenza sessuale

Ognuna di queste misure era valutata in 3 studi diversi.

Tabella 2A. Misure di struttura per il cancro alla prostata utilizzate in letteratura

Misura	Usata da	Tipo di dati	Sample size	Istituzione
Proporzione consistente di pazienti trattati da provider ad alti volumi (terzile più alto) – riferito al chirurgo e al radioterapista	{Shroeck F. R. et al. (2014)}	Survey	48050	Accademica, non-profit, for-profit
	{Herrel L. A. et al. (2016)}	Flussi amministrativi	72411	Accademica, non-profit, for-profit
	{Schroeck F. R. et al. (2015)}	EHRs	38055	Accademica, non-profit, for-profit
	{Schroeck F. R. et al. (2014 April)}	EHRs	53614	Accademica, non-profit, for-profit
	{Schroeck F. R. et al. (2014 Nov)}	EHRs	462747	Accademica, non-profit, for-profit
	{Miller D.C. et al. (2007)}	EHRs	5230	Di insegnamento, non-profit
Disponibilità di strutture di oncologia e radiologia e la presenza di una consulenza psicologica per i pazienti	{Miller D.C. et al. (2007)}	EHRs	5230	Di insegnamento, non-profit
Presenza nella struttura di una certificazione del board degli urologi e dei radiologi oncologi	{Bekelman J.E. et al. (2007)}	EHRs e flussi amministrativi	11674	Di insegnamento, non-profit e for-profit
	{Miller D.C. et al. (2007)}	EHRs	5230	Di insegnamento, non-profit

Tabella 2B. Misure di processo per il cancro alla prostata utilizzate in letteratura

Misure	Usato da N studi	Data Source	Type of institution
Terapia di deprivazione androgenica (ADT)/terapia ormonale adiuvante per pazienti con tumore prostatico ad alto rischio	15	Survey, EHRs, e dati amministrativi	Accademica, non-profit, for-profit Federale
Documentazione clinica dello stadio T	10	Survey, EHRs, dati amministrativi	Accademica, non-profit, for-profit
Evitare misure eccessive - Scansione ossea per la stadiazione di pazienti a basso rischio	10	Survey, EHRs, dati amministrativi	Accademica, for-profit, federale
Documentazione dei livelli del PSA (Prostate Specific Antigen)	9	EHRs, Survey, dati amministrativi	Accademica, non-profit, for-profit
Reporting dei livelli di PSA al paziente	9	Survey, EHRs, dati amministrativi	Accademica, for-profit, federale
Counselling multidisciplinare sulle opzioni di trattamento	9	Survey, EHRs, dati amministrativi	Accademica, non-profit, for-profit
Seguire il paziente almeno 2 volte nel primo anno di trattamento	7	Survey, EHRs, dati amministrativi	Accademica, non-profit, for-profit
Valutazione delle comorbidità	6	Survey, EHRs, dati amministrativi	Accademica, for-profit

Reporting di patologie conseguenti alla PR	4	Survey, EHRs, dati amministrativi	Accademica, non-profit, for-profit
Valutare lo stadio della malattia prima della stadiazione clinica trattamento o pre-trattamento attraverso DRE, PSA totale, Gleason Grade	3	Survey, EHRs, dati amministrativi	Accademica, for-profit, federale (VA)
Valutazione della storia familiare per Ca prostatico	3	Survey, EHRs	Insegnamento non-profit
Scansione ossea eseguita in pazienti ad alto rischio	3	EHRs, dati amministrativi	Accademica, non-profit, for-profit
Evitare misure eccessive - TAC per la stadiazione di pazienti a basso rischio	3	Survey, EHRs, dati amministrativi	Accademica, non-profit, for-profit
Dosi sull'asse centrale di almeno 75 Gy durante la radioterapia	3	EHRs, dati amministrativi	Federale, di insegnamento, non-profit

Tabella 2C. Misure di esito per il cancro alla prostata utilizzate in letteratura

Measures	Used By	Data	Sample	Type of institution
Ricovero in ospedale o trattamento medico o chirurgico per una serie di gravi complicanze	Worwag et al. (1998)	Survey	100	Ospedale universitario, non profit
Evitare ricoveri ospedalieri multipli negli ultimi 30 giorni di vita	Herrel I. A. et al. (2016)	Administrative	72411	Accademica, non-profit, for-profit
Ammissione all'Hospice negli ultimi giorni di vita	O'Connor N. R. et al. (2014)	EHRs	2929	Non-profit (hospice)
Margini chirurgici ben conservati	Webster T. M. et al. (2012)	EHRs	133	non-profit (comunità)
Valutazione del funzionamento urinario, sessuale e intestinale del paziente dopo il trattamento	Hernandez-Boussard et al. (2016)	EHRs	5349	Ospedale universitario
	Shroeck F. R. et al. (2015)	EHRs	38055	Accademica, non-profit, for-profit
	Wei J. T. et al. (2002)	Survey	1213	Ospedale di insegnamento
Soddisfazione dei pazienti per la scelta del trattamento, per la continenza e la potenza sessuale	Kamal A. H. et al. (2013)	Survey	459	Non-profit (di comunità)

	Worwag et al. (1998)	Survey	100	Accademia, ospedale, non-profit
	Wright J. D. et al. (2015)	Survey	14772	Accademia, non-profit, for-profit

La Tabella 3 illustra il ranking delle prime dieci misure di qualità in base al numero di studi che le hanno utilizzate. La maggior parte, abbiamo visto, erano misure di processo. La deprivazione androgenica/terapia ormonale adiuvante per pazienti ad alto rischio si è classificata come la prima misura valutata in letteratura, con ben 15 studi che la citavano come misura di qualità. La ricerca era lo scopo principale della maggior parte di questi studi, come evidenziato anche dalla Figura 2, con un solo studio inteso a migliorare la qualità effettiva.

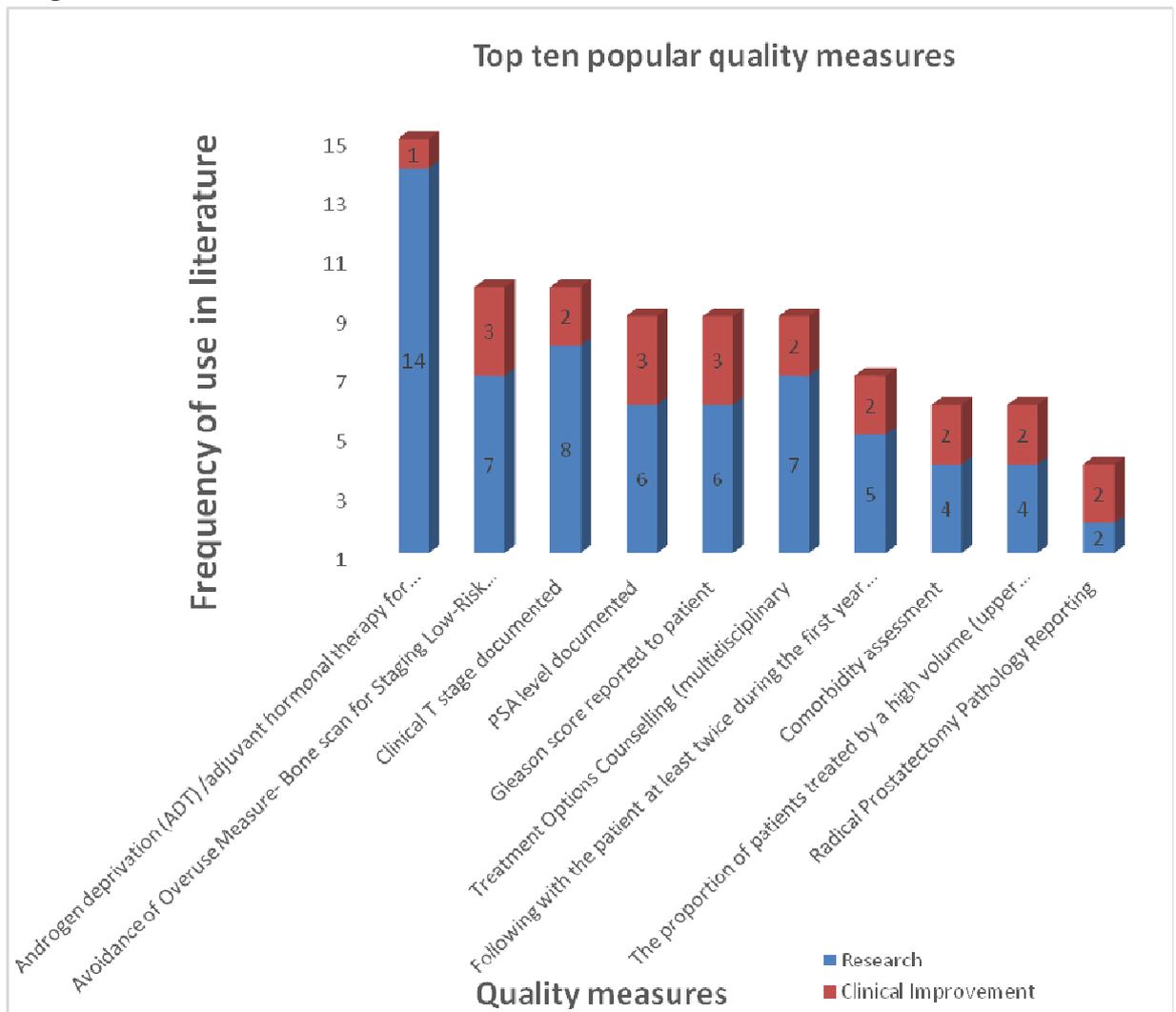
Evitare l'uso eccessivo di scansioni TAC per la stadiazione di pazienti a basso rischio; e la documentazione clinica della fase T si sono entrambe classificate al secondo posto con 10 studi ognuna. Dei 10 studi associati alla prevenzione dell'uso eccessivo della TAC, 7 studi hanno riguardato la ricerca e 3 studi erano invece correlati ad un effettivo miglioramento della qualità (Figura 2). Allo stesso modo, la ricerca era lo scopo di 8 studi e il miglioramento della qualità in 2 studi che citavano la presenza di documentazione clinica dello stadio T (Figura 2).

In generale, la ricerca era lo scopo prevalente di tutte le prime dieci misure considerate (Figura 2).

Tabella 3. “Top-ten” delle misure di qualità per il cancro alla prostata più usate

Rank	Misure	Numero di studi
1	Terapia di deprivazione androgenica (ADT)/terapia ormonale adiuvante per pazienti con tumore prostatico ad alto rischio	15
2	Documentazione clinica dello stadio T	10
2	Evitare misure eccessive - Scansione ossea per la stadiazione di pazienti a basso rischio	10
3	Documentazione dei livelli del PSA (Prostate Specific Antigen)	9
3	Reporting dei livelli di PSA al paziente	9
3	Counselling multidisciplinare sulle opzioni di trattamento	9
4	Seguire il paziente almeno 2 volte nel primo anno di trattamento	7
5	Valutazione delle comorbidità	6
5	Proporzione consistente di pazienti trattati da provider ad alti volumi (terzile più alto) – riferito al chirurgo e al radioterpista	6
6	Reporting di patologie conseguenti alla PR	4

Figura 2: Distribuzione stratificata delle misure “top-ten”, a seconda che gli studi fossero focalizzati alla ricerca o al miglioramento clinico.



Lo studio sul miglioramento della qualità era direttamente correlato all'aumento della qualità dell'assistenza stessa. Flynn et. al. (Flynn, 2010), per esempio, intervistavano i medici di medicina generale praticanti in due stati nord-americani. La valutazione della storia familiare come misura di qualità si è basata sull'estensione della storia familiare e sull'età della diagnosi di cancro per i familiari colpiti.¹⁵ Lo studio, d'altra parte, potrebbe non essere direttamente associato al miglioramento della qualità. Ad esempio, Bekelman et al. (Bekelman, 2007) hanno condotto uno studio per valutare le variazioni di aderenza alle misure di qualità della radioterapia con external beam (EBRT) per la cura localizzata del cancro alla prostata.

I ricercatori hanno quindi valutato l'aderenza a cinque misure di qualità EBRT, inclusa la valutazione del completamento di due visite di follow-up con un oncologo radioterapista nell'anno successivo alla terapia utilizzando i dati di registro Surveillance Epidemiology, and End Result (SEER). Il Gleason score, la documentazione dei livelli di PSA e il counselling sulle opzioni di erano al terzo posto, ognuna con 9 studi.

Nel Gleason Score e nella documentazione dei livelli di PSA 6 studi erano associati alla ricerca e 3 misuravano miglioramenti effettivi nella qualità. Dei nove studi sulle opzioni di trattamento 7 erano per la ricerca e due per il miglioramento della qualità.

Seguire il paziente con almeno due visite nell'anno successivo al trattamento era al quarto posto, con 5 studi di ricerca e 2 di miglioramento della qualità.

In modo simile, la valutazione delle comorbidità e la proporzione di pazienti trattati da provider ad alti volumi erano al quinto posto, con 6 studi ognuno. Di questi la divisione fra ricerca e applicazione alla qualità era sempre di 4 a 2.

Infine, il reporting della anatomia patologica dopo prostatectomia radicale era al sesto posto, con due studi di ricerca e due di miglioramento della qualità

Discussione

Tramite questo studio, in letteratura, abbiamo identificato 71 misure di qualità proposte per valutare il cancro alla prostata. Le misure di qualità più frequentemente riportate sono state associate ai processi di cura (misure di qualità del processo), e tra le misure meno studiate vi erano quelle relative a importanti esiti dei pazienti, come la valutazione della funzione sessuale e dell'incontinenza urinaria. Tra tutti i parametri di qualità identificati, solo 3 sono stati segnalati in più di 10 studi (vale a dire documentazione sulla fase clinica, evitare la TAC per la stadiazione di pazienti a basso rischio e uso di terapia ormonale adiuvante per pazienti ad alto rischio). In generale la maggior parte era usata solo in uno o due studi. Dato il numero abbondante di misure di qualità proposte per valutare il cancro alla prostata e il loro utilizzo limitato, abbiamo sintetizzato queste informazioni come base per lo sviluppo di una serie di future misure di qualità proposte per valutare il cancro alla prostata che potrebbero essere al centro di sforzi di ricerca e miglioramento della qualità.

Tra le dieci misure di qualità più comunemente utilizzate in letteratura, nove sono state le misure di qualità di processo, che misuravano una valutazione pre-trattamento trattamento o un ordine eseguito dal team di assistenza. Altri studi dimostravano anche che la maggior parte delle metriche utilizzate nei programmi governativi si concentravano solo su misure di processo, molte delle quali non erano clinicamente significative. (Greenhalg, 2007) È estremamente probabile che l'abbondanza di misure di processo sia dovuta al fatto che esse sono generalmente più facili da misurare e implementare. I risultati di questo studio hanno mostrato che la maggior parte degli studi che utilizzavano queste misure di qualità del processo erano focalizzati sulla

ricerca, con solo un numero limitato di studi che poi ne riportavano l'uso per il miglioramento della qualità, e quindi ancora una volta sottolineandone la mancanza di prove di efficienza. Ciò indica un divario tra la ricerca accademica e la pratica clinica reale nel migliorare le prestazioni su queste misure di qualità, almeno in termini di ciò che è disponibile nella letteratura medica. L'integrazione della ricerca e della pratica clinica per il miglioramento della qualità potrà sicuramente in futuro aiutare a ridurre il divario esistente, e portare all'implementazione di misure davvero trasferibili nel contesto della pratica.

Dato ancora più importante ai fini di questa tesi, abbiamo anche scoperto che solo una piccolissima parte delle misure proposte erano misure di esito e che esse vengono peraltro utilizzate raramente, sia per la ricerca che per il miglioramento della qualità. Nella letteratura recente, Jha et al. hanno sottolineato quanto sia importante l'eliminazione di parametri di qualità non necessari dai programmi governativi (ad esempio, la remunerazione per le prestazioni o come anzidetto molti degli indicatori di processo, talvolta ridondanti), per focalizzarsi verso progressi significativi sulla sicurezza dei pazienti e la loro qualità di vita post intervento di PR. (Jha 2013, Jha 2016)

La relativa mancanza di misure di outcome proposte e utilizzate per il cancro alla prostata identifica una disconnessione tra i dati disponibili alla misurazione e le misure di qualità che davvero contano sia per i providers che, soprattutto, per i pazienti. Come discusso in una delle recenti bozze del NQF americano, è quindi fondamentale ottenere sempre il feedback dai pazienti e dalle loro famiglie nell'attuazione di qualsiasi misura. (Report NQF, 2017) Il mancato sviluppo e utilizzo di misure di outcome può rappresentare quindi un limite allo sviluppo di futuri programmi nazionali o regionali. Tale limite andrà quindi, affrontato e superato, in modo da misurare indicatori che siano in grado di coinvolgere in primis i clinici per la misurazione della qualità, e, non meno importante, a coinvolgere i pazienti

con la valutazione di esiti davvero impattanti nella loro esperienza di malattia e nella relativa qualità di vita percepita.

Infine, il nostro studio oltre che enucleare una serie di indicatori “comunemente misurati” – la Top ten-, ma ha individuato 5 indicatori fondamentali (i più utilizzati e considerati dalla letteratura maggiormente impattanti) (Tabella 4) per ciascuna delle tre classificazioni teorizzate da Donabedian. In particolare, per le misure di esito, sono stati selezionati due indicatori:

- *“Valutazione del funzionamento urinario, sessuale e intestinale del paziente dopo il trattamento”;*
- *“Soddisfazione dei pazienti per la scelta del trattamento, per la continenza e la potenza sessuale”*

Il primo di essi, nella parte sperimentale successiva di questa tesi sarà misurato e confrontato con il reporting dei clinici in un caso di studio concreto di un ospedale universitario americano – Stanford, California.

Tabella 4 Proposta di una serie di misure di qualità essenziali emerse dalla letteratura scientifica

Misure	Tipo
La proporzione di pazienti trattati da un provider ad alti volumi	Struttura
Terapia di deprivazione androgenica (ADT)/terapia ormonale adiuvante per pazienti con tumore prostatico ad alto rischio	Processo
Documentazione clinica dello stadio T	Processo
Valutazione del funzionamento urinario, sessuale e intestinale del paziente dopo il trattamento	Esito
Soddisfazione dei pazienti per la scelta del trattamento, per la continenza e la potenza sessuale	Esito

Caso di studio: potenzialità per la raccolta dei dati PROMs dalle cartelle cliniche

I risultati e la discussione del nostro studio di revisione sottolineano una relativa mancanza, in letteratura, di misure di esito, con una prevalenza e una preferenza netta, da parte degli studi scientifici, dell'analisi di processo. Se da un lato non è scopo di questa tesi criticare l'utilità della misurazione di queste metriche (in particolare per il loro effetto sulle decisioni organizzative e gestionali dei processi di diagnosi e cura), dall'altro lato l'analisi precedente ha fatto emergere come vi sia una relativa mancanza di misure di esito. Come detto in precedenza, le misure di esito e di esperienza, specialmente se riportate sui pazienti (PREMs e PROMs) riescono a meglio indirizzare decisioni e pratiche che possono avere un impatto maggiore sulla qualità della vita dei pazienti.

L'incontinenza urinaria (IU) e la disfunzione erettile (DE) sono gli effetti collaterali più frequenti della PR (Barocas 2017, Donovan 2016) e quelli che influenzano maggiormente la qualità della vita post intervento (Chen 2017, Penson 2005, Stanford 2000, Litwin 1995). Questi outcome centrati sul paziente (PCO) sono principalmente riportati dai pazienti (PROMs), ma i tassi di questi effetti collaterali sono apparentemente molto più elevati rispetto a quanto riportato dai clinici e talvolta difformi fra le diverse popolazioni studiate, dipendendo soprattutto dalle caratteristiche del campione e/o del trattamento (Whiting 2016, Sanda 2008 e Miller 2005).

Una valutazione accurata degli esiti centrati e rilevati dal paziente dopo la terapia per il carcinoma prostatico localizzato è fondamentale nel consigliare i pazienti per quanto riguarda la selezione del trattamento più efficace per il miglioramento degli esiti del trattamento. Sfortunatamente, i dati relativi ai risultati dopo il trattamento del carcinoma prostatico localizzato sono limitati a serie di casi provenienti da singole istituzioni, (Walsh, 2011) survey di pazienti che usano strumenti standardizzati, (Barocas, 2017; Chen 2017) o studi osservazionali basati sui registri. (Kowalczyk, 2012) Di questi metodi,

solo le indagini sui pazienti forniscono stime accurate dei PROMs, anche se l'implementazione di survey può essere costosa e di difficile applicazione per pazienti e medici. Inoltre, in assenza di questi dati, è molto difficile per i medici ottenere il feedback necessario per migliorare i risultati della chirurgia. Di conseguenza, i recenti sforzi per stabilire linee guida per il trattamento del cancro alla prostata sono stati inconcludenti a causa di prove insufficienti in merito ai benefici e ai rischi relativi delle diverse opzioni di trattamento possibili.

L'adozione diffusa di cartelle cliniche elettroniche (EHR) offre l'opportunità di estrarre informazioni sugli esiti correlati al trattamento dalle note cliniche regolarmente registrate come testo libero nella EHR. L'avvento di strumenti sempre più sofisticati, come l'elaborazione del natural language processing (NLP), può essere utilizzato per identificare ed estrarre in modo efficiente i dati non strutturati dalle EHR, come i risultati riportati dai medici. (Hernandez-Boussard, 2016) Tuttavia, una preoccupazione significativa è che i medici non riportino degli accurati PROMs nelle loro note. (Basch, 2010, Flores, 2012)

Sulla base di dati precedenti che mostrano la concomitanza di esiti registrati dal medico e quelli da indagini oggettive in uomini sottoposti a prostatectomia radicale, abbiamo ipotizzato che i risultati registrati nelle note del medico si correlassero bene con quelli riportati da pazienti che utilizzavano strumenti di risultato come l'indice Expanded Prostate Cancer Index Composite (EPIC). Per verificare questa ipotesi, abbiamo effettuato un processo di *data-mining* nelle EHR per estrarre i risultati riportati dal medico per l'incontinenza urinaria dopo la prostatectomia radicale e confrontarli direttamente con quelli riportati utilizzando lo strumento validato EPIC per valutare i risultati urinari nei pazienti con cancro alla prostata. (Wei, 2000, Szymansky 2010)

Metodi

Il questionario EPIC 26

L'Expanded Prostate Cancer Index Composite (EPIC) –già nominato precedentemente fra i *gold standard* utilizzati per la valutazione del carcinoma prostatico nell'ambito oncologico - è un questionario validato di 50 elementi liberamente disponibile dal 2000 (Wei, 2000) che affronta i domini: disfunzioni urinarie, intestinali, sessuali e ormonali sulla qualità della vita. Nel 2010 è stato sviluppato un questionario più breve di 26 punti – EPIC-26 (Szymanski, 2010). Questo questionario è stato validato in diverse lingue (non in lingua italiana) ed è uno dei questionari più utilizzati in tutto il mondo per valutare la qualità della vita dopo prostatectomia radicale (Lee, 2015).

Sorgente dei dati

Abbiamo utilizzato i dati del database Stanford University HealthCare (SHC) (2005-2017) accessibili tramite un protocollo approvato dal comitato etico dello Stanford Human Subjects Research. Il sistema informatico fornisce dati sui servizi ospedalieri, ambulatoriali e di cure primarie. I dati dei pazienti includevano dati demografici, accessi, elenchi di diagnosi/problemi, report clinici, note sanitarie/documentazione delle visite, risultati di laboratorio e diagnostici, farmaci, piani di trattamento, procedure, anamnesi del paziente, indagini sui pazienti e informazioni di follow-up/sopravvivenza.

Abbiamo effettuato una procedura di estrazione con *data-mining* per analizzare la coorte di pazienti affetti da cancro alla prostata, questa procedura era stata precedentemente validata e implementata. (Hernandez-Boussard, 2016; Hernandez-Boussard, 2017) In breve, essa funziona tramite la cattura di tutti i codici di diagnosi ICD-9/ICD-10 delle procedure ICD-9 e dei codici del Current Procedural Terminology con la cattura anche degli elenchi di farmaci, dei valori di laboratorio, dei test diagnostici, degli

incontri/visite col paziente, dei farmaci, dei piani di trattamento e, infine, di tutte le note cliniche associate per tutti i pazienti nella coorte. Gli identificativi dei pazienti sono stati collegati al California Cancer Registry (CCR) per ottenere informazioni anche sui trattamenti al di fuori dell'università di Stanford. L'analisi istologica e le informazioni sulla patologia sono state usate al fine di calcolare con precisione Grado e Stadio. Le note cliniche e il testo narrativo sono stati analizzati per identificare i termini di interesse utilizzando un software open source basato su Java GATE (General Architecture for Text Engineering). (Heffner, 2009) Abbiamo usato le espressioni regolari di ConText per contrassegnare i termini negati (ad esempio "nessuna lamentela di incontinenza urinaria") e menzioni di eventi familiari o anamnesi recente e remota. (Harkema, 2009) I risultati sono stati archiviati in un database Oracle che include dati demografici dei pazienti, caratteristiche cliniche, nonché menzioni affermate e negate dei termini clinici trovati nel testo e per noi rilevanti, che sono stati infine convertiti in una matrice per effettuare l'analisi.

Popolazione in studio

Il cancro alla prostata è stato identificato nelle EHR utilizzando codici ICD-9 / ICD-10, codici CTP o CCR. I pazienti affetti da cancro alla prostata sottoposti a prostatectomia radicale sono stati reclutati per partecipare a una valutazione PROMs prospettica - valutata con questionario EPIC-26 al basale (prima della prostatectomia) e a 3, 6, 12 e 24 mesi dopo la prostatectomia. Le survey effettuate dai pazienti sono state collegate alla coorte di pazienti con EHR. I medici sono stati messi in cieco rispetto alle risposte date dai pazienti al questionario EPIC al momento della visita.

Esiti misurati e covariate prese in considerazione per questa analisi

Per estrarre le menzioni di “fastidio urinario” dagli EHR, un infermiere del reparto di urologia ha riesaminato manualmente 100 cartelle selezionate in modo random per sviluppare un vocabolario di estrazione sulla base dei termini più correlati a “incontinenza urinaria” e “sintomi di una minzione fastidiosa e/o irritativa e difficoltà allo svuotamento della vescica”. Quattro clinici di diversi istituti hanno convalidato i termini clinici. Un fenotipo elettronico è stato progettato per identificare ed estrarre valutazioni cliniche sulla base degli EHR. I fenotipi includono codici ICD-9 / ICD-10, codici di rimborso, farmaci e parole abbinati a ontologie esistenti e fornite dal National Center for Biomedical Ontology (NCBO) e dall’ Unified Medical Language System (UMLS). (Muse, 2012) Utilizzando la nostra procedura per effettuare il *data-mining*, abbiamo identificato ed estratto affermazioni cliniche positive e negative di incontinenza urinaria/disturbo dagli EHR e basate quindi sul fenotipo elettronico. Il fenotipo da noi sviluppato è disponibile attraverso il repository pubblicamente accessibile PheKB. (Kyrby,2016)

In EPIC-26, i punteggi per il disturbo urinario sono stati calcolati utilizzando i metodi di punteggio pubblicati precedentemente. (Wei, 2002) Per avere un confronto diretto tra questionario EPIC-26 ed EHR, abbiamo creato una variabile dicotomica dalle domande EPIC-26 che valutano l’"Incontinenza urinaria".

Esse sono:

Incontinenza urinaria (q23: *"Nelle ultime 4 settimane, quanto spesso ha manifestato perdite di urina?"*)

Controllo vescicale (q26: *"Quale delle seguenti descrizioni meglio indica il suo controllo urinario durante le ultime 4 settimane?"*)

Queste due variabili sono state utilizzate come variabili di prova (un "gold standard" appositamente creato).

In dettaglio, usando operatori booleani per una migliore spiegazione, abbiamo classificato un paziente come:

- "*continente*" se ha risposto: 4 ("Circa una volta alla settimana") OPPURE 5 ("Raramente o mai") alla q23 E ha risposto 3 ("Dribbling occasionale") OPPURE 4 ("Controllo totale") a q26

- "*incontinente*" se ha risposto: 1 ("Più di una volta al giorno") OPPURE 2 ("Circa una volta al giorno") OPPURE 3 ("Più di una volta alla settimana") a q23 E 1 ("Nessun controllo urinario in assoluto") OPPURE 2 ("Dribbling frequenti") OPPURE 3 ("Dribbling occasionali") a q26

Se mancava anche solo una delle domande suddette, il paziente non era classificato. Se il paziente rispondeva in modo discordante alla classificazione precedente, anche in questo caso era considerato come mancante.

L'incontinenza urinaria valutata sulla base di queste domande è stata confrontata per un testing di accuratezza con l'uso del pannolino (q27: "*Lei utilizza pannolini?*" - variabile dicotomica a cui è possibile rispondere "sì" o "no") e valutazione generale del disturbo/fastidio (q28: "*Nel complesso, quanto è stato importante il suo problema nella funzione urinaria nelle ultime 4 settimane?*" - In questo caso abbiamo creato una variabile dicotomica unendo le risposte "Nessun problema o piccolo problema" vs "problema moderato o grande").

I dati demografici e clinici sono stati estratti dalle EHR. Il punteggio dell'indice di comorbidità di Charlson (CCI) è stato utilizzato per definire la gravità della comorbidità e classificato in tre livelli: lieve (1-2); moderato (3-4); e grave (≥ 5). (D'Hoore, 1993)

Analisi statistica

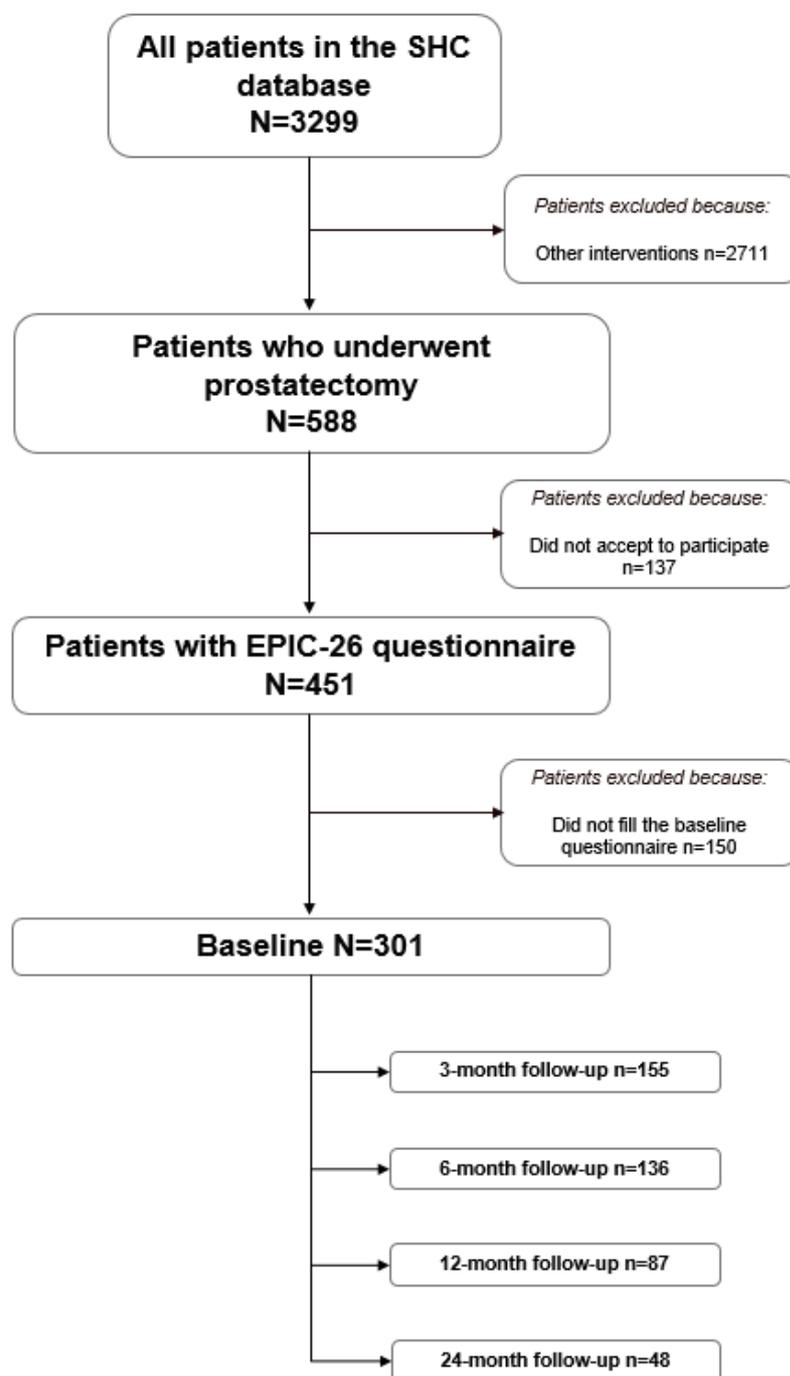
Le variabili continue sono state riassunte come media \pm deviazione standard (SD), le variabili categoriali come frequenze e percentuali. Il test T di Student è stato utilizzato per confrontare i valori medi tra i gruppi. La prevalenza di IU ad ogni punto temporale, valutata mediante EHR ed EPIC-26, è stata confrontata usando ANOVA su misure ripetute. L'accordo della IU riportata tra EHR ed EPIC-26 è stato valutato utilizzando il kappa di Cohen in ciascun momento e valutato come discreto, moderato, sostanziale e quasi perfetto utilizzando i valori limite di 0,21, 0,41, 0,61 e 0,81 definiti da Landis e Koch. (Landis, 1977)

Sono stati utilizzati modelli di regressione logistica univariata per testare la potenziale associazione di variabili demografiche e cliniche con il disaccordo nel reporting di IU fra EPIC-26 ed EHR al tempo basale e 3 mesi. Le variabili testate includevano età, etnia, razza, anno di diagnosi, indice di massa corporea, stato civile, lingua, stato assicurativo, punteggio di Charlson, stadio e grado di cancro. Un valore $p < 0,05$ è stato considerato statisticamente significativo. Tutte le analisi statistiche sono state eseguite utilizzando STATA, versione 14 e SPSS, versione 21.22,23 I calcoli di potenza post hoc sono stati effettuati utilizzando il software statistico G-Power, versione 3.1.9.2.

Risultati

Un totale di 451 soggetti ha acconsentito a partecipare a questo studio e 301 hanno completato almeno la valutazione di base EPIC-26 e hanno avuto almeno una valutazione dell'incontinenza urinaria nella cartella clinica al basale o durante il follow-up (Figura 3).

Figura 3: Flowchart per la selezione della coorte



Le caratteristiche demografiche e cliniche di questo sottogruppo sono state confrontate con quelle di coloro che non hanno completato la valutazione di base (N = 150) e non sono state trovate differenze significative, suggerendo che i pazienti che hanno completato l'EPIC-26 erano rappresentativi della coorte complessiva sottoposta a intervento chirurgico.

Le caratteristiche demografiche e cliniche di base dei 301 partecipanti sono visualizzate nella Tabella 5.

Tabella 5: caratteristiche cliniche e demografiche al baseline

Caratteristiche demografiche	Media (\pmDS) o N(%)
Eta, media (DS), y	61.4 (\pm 6.5)
Etnia, N	
Binca/caucasica	228 (75.7)
Nera/africana	8 (2.7)
Asiatica	16 (5.3)
Ispanica	8 (2.7)
Isole del pacifico	1 (0.3)
Altri/sconosciuta	40 13.3
Livello di istruzione, N=301	
< Scuole superiori	8 (2.7)
Scuole superiori	25 (8.4)
Parte di percorso universitario	70 (23.2)
Laurea	78 (26)

Formazione post laurea	120 (39.7)
Stato civile (n=291)	
Sposato/partner per la vita	247 (84.9)
Divorziato/single/separato/altro	44 (15.1)
Parlante inglese (n=288)	281 (97.5)
Stato assicurativo (n=296)	
Medicare	119 (39.5)
Medicaid	2 (0.7)
Privato	173 (57.5)
Altro	2 (0.7)
Anno di diagnosi	
2005-2010	145 (48.2)
2011-2017	156 (51.8)
Caratteristiche cliniche	
Terapia adiuvante	
RT	19 (6.3)
Terapia ormonale	16 (5.3)
Grado (n=295)	
1	12 (4.1)
2	104 (35.3)
3	176 (59.7)

4	3 (1.0)
N/A	6 (2.0)
Stadio (n=297)	
1	22 (7.3)
2	194 (64.5)
3	67 (22.3)
4	14 (4.7)
Comorbidità	
Charlson's score, media (DS)	2.98 (\pm 1.95)

L'età media alla chirurgia era di $61,4 \pm 6,5$ anni e il 76% era di etnia Caucasica. Circa il 40% aveva un alto livello di istruzione e l'85% era sposato. L'indice di gravità di Charlson sulle comorbidità aveva una media di 2,98 punti, e una deviazione standard di 1,95, che denotava una moderata gravità data dalle comorbidità.

La sensibilità e la specificità della misurazione della IU tramite EPIC-26 (calcolata rispetto al nostro "gold standard") erano superiori al 95% e all'85% rispettivamente in tutti i punti (Tabella 6).

Tabella 6: Sensibilità e specificità della definizione di IU ai diversi punti*

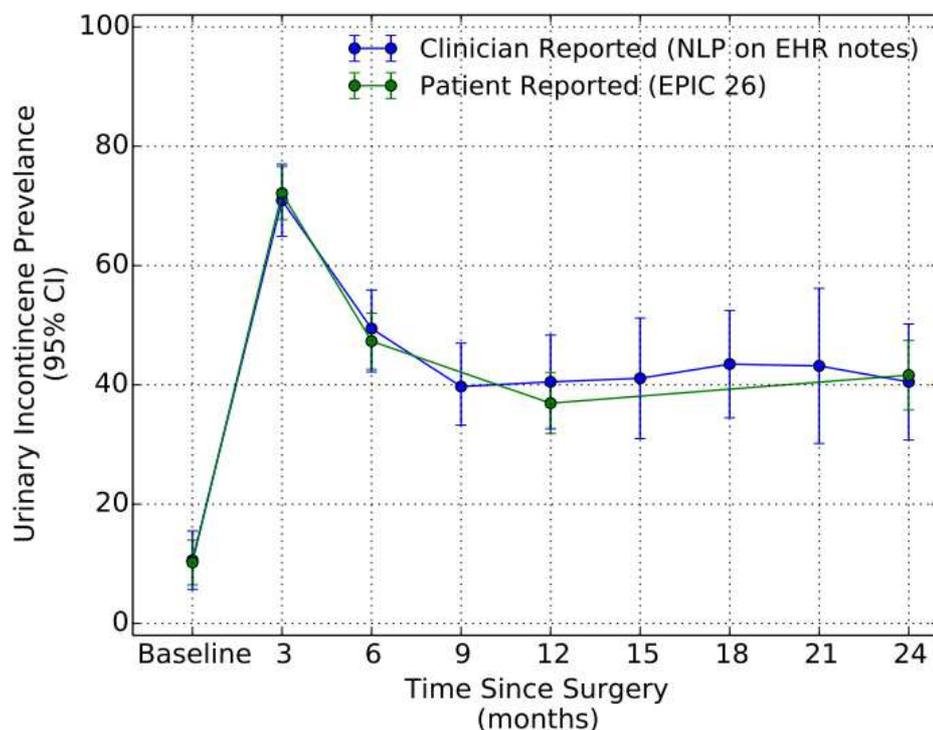
Definizione di IU	Sensibilità**	Specificità**
Baseline	97.53% (95.73 – 98.72)	100% (99.19 – 100)
3 mesi	97.44% (91.04 – 99.69)	86.27% (78.04 – 92.29)
6 mesi	96.43% (91.86 – 98.83)	86.11% (80.18 – 90.81)
12 mesi	100% (97.63 – 100)	94.25% (89.68 – 97.21)
24 mesi	95.18% (89.62 – 98.88)	95.11% (91.22 – 98.53)

*IU (come dicotomica) nella q23 (incontinenza) e q26 (controllo della vescica) come comparata alla q27 (uso del pannolino) e q28 (fastidio urinario)

** Percentuale e 95% IC

La documentazione dei clinici per quanto riguardava la IU dei pazienti nelle EHR è stata confrontata con quella ottenuta direttamente dal paziente attraverso l'indagine EPIC-26. Ciò è stato fatto sia al basale che ai successivi follow-up. Il momento del completamento del questionario EPIC e la visita clinica erano uguali in tutti i casi, dal momento che il questionario è stato raccolto poco dopo la visita clinica. Nel corso di 24 mesi di follow-up postoperatorio, le percentuali di IU estratte dalle note cliniche nelle EHR erano molto simili a quelle riportate direttamente dai pazienti su EPIC-26 (Figura 4).

Figura 4: prevalenza di IU dalle cartelle cliniche e da EPIC-26 al baseline e ai vari punti del follow-up



La prevalenza di IU al basale (preoperatoria) è stata riportata dal 10% dei pazienti nel questionario EPIC 26 (Figura 4), e questo tasso di incontinenza era molto simile a quanto registrato dai medici sulla EHR (10,6%, $p = 0,90$). L'IU riportata su EPIC-26 è aumentata al 72,1% entro 3 mesi dall'intervento e scesa al 47,3% a 6 mesi, con percentuali quasi identiche riportate sia dai medici nella cartella clinica (70,9%, $p = 0,75$ a 3 mesi e 49,4%, $p = 0,60$ a 6 mesi). A 12 e 24 mesi, i tassi sono scesi a circa il 40% sia nel questionario EHR sia in quello EPIC-26 ($p = 0,45$ e $0,85$, rispettivamente). Inoltre, abbiamo calcolato i punteggi UI di EPIC-26 nella nostra popolazione di studio al basale e nei successivi follow-up e abbiamo verificato come fossero molto simili a quelli riportati in diversi studi recenti (Tabella 7).

Tabella 7: Score di IU da EPIC-26 nei differenti punti e comparazione con studi condotti in precedenza.

Tempo	Questo studio N=301	Donovan et al. 2016 N=553	Chen et al. 2017 N=462	Barocas et al. 2017 N=1523	p
Baseline	92.8 (1.1)	92.8 (3.4)	90.3 (2.5)	86.7 (1.2)	n.s.
3 mesi	52.1 (2.9)	-	54.8 (5.3) – 2 mesi	-	n.s.
6 mesi	68.9 (2.7)	67.4 (5.3)	-	65.4 (1.6)	n.s.
12 mesi	74.5 (2.9)	76.5 (4.9)	67.7 (5.4)	70.7 (1.5)	n.s.
24 mesi	75.2 (2.9)	80.3 (4.6)	67 (4.8)	71.8 (1.4) - 3 anni	n.s.

I valori di Kappa tra EPIC-26 ed EHR hanno mostrato un moderato accordo al basale, 3 mesi, 6 mesi e 24 mesi e accordo sostanziale a 12 mesi di follow-up (Tabella 8), indicando come i valori simili riportati dai pazienti o estratti dalle EHR fossero molto difficilmente imputabili alla semplice casualità. Sono stati utilizzati modelli di regressione logistica per valutare se i fattori clinici e demografici rappresentassero delle differenze nella IU estratta da EHR e quella riportata da EPIC-26 (Tabella 9). Come visto nella tabella, non c'erano significativi fattori predittivi di disaccordo tra IU riferita dal paziente e quella trascritta dal medico su EHR al basale. È interessante notare che, a 3 mesi di follow-up, l'età era un importante predittore di disaccordo nel riportare IU (OR: 0,93, IC 95% 0,87 - 0,99, $p = 0,032$) con uomini più giovani che mostravano una maggiore probabilità di disaccordo rispetto ai medici, e un maggiore reporting del sintomo IU su EPIC-26. Quando il campione dei pazienti era dicotomizzato all'età di 60 anni, la probabilità di disaccordo era del 66% più alta per i pazienti di età <60 anni (OR: 0,34, IC 95% 0,15 - 0,77, $p = 0,01$). Nello specifico, i medici hanno riportato in forma minore l'IU in pazienti più giovani 8/36 (22%) (<60 anni di età) rispetto a pazienti più anziani 9/76 (11,8%) (età ≥ 61 anni). (Tabella 10)

Tabella 8: Percentuale di accordo osservato tra EPIC-26 e cartelle cliniche, accordo atteso per effetto casuale da solo e Cohen's Kappa.

Incontinenza Urinaria	Accordo osservato	Accordo atteso	Livello di agreement	Cohen's Kappa	Errore std.	p
Baseline	90.7%	81%	Moderato	0.51	0.06	<0.0001
3 mesi	80%	59%	Moderato	0.51	0.08	0.0001
6 mesi	71.3%	50%	Moderato	0.41	0.09	0.0001
12 mesi	83.9%	50.8%	Sostanziale	0.67	0.10	0.0001
24 mesi	75%	50%	Moderato	0.50	0.14	0.0002

Kappa <0 indica nessun accordo, 0–0.20 leggero, 0.21–0.40 discreto, 0.41–0.60 moderato, 0.61–0.80 sostanziale, e 0.81–1 come accordo quasi perfetto.

Tabella 9. Predittori di disaccordo sulla presenza di IU tra EPIC e EHR. Risultati dal modello di regressione logistica univariata.

	Baseline			3-mesi		
	OR	95% IC	p	OR	95% IC	p
Anno (≥ 2011 vs. < 2011)	1.08	0.49-2.36	0.85	0.72	0.32-1.63	0.43
BMI	0.97	0.85-1.08	0.58	1.08	0.97-1.20	0.14
Età (≥ 60 vs. < 60)	0.75	0.34-1.64	0.47	0.34	0.15-0.77	0.01
Grado	1.36	0.66-2.77	0.40	0.79	0.43-1.46	0.45
Stadio	1.11	0.62-1.98	0.73	1.04	0.59-1.84	0.90
Etnia						
Ispanica	0		0.99	0		0.99
sconosciuta	0.77	0.17-3.46	0.73	0.98	0.26-3.73	0.98
Razza (bianca vs. altre)	0.94	0.38-2.30	0.88	0.73	0.29-1.85	0.51
Stato civile	1.03	0.34-3.13	0.96	2.06	0.45-9.50	0.36
Lingua	0.61	0.07-5.28	0.66	0.74	0.07-7.35	0.80
Tipo di assicurazione (MediCare/ MedicAID vs. privato/altro)	0.69	0.27-1.78	0.44	0.67	0.30-1.50	0.33
Indice di Charlson						
3-4 vs. 1-2	2.41	0.85-6.83	0.10	0.98	0.29-3.34	
>4 vs. 1-2	1.91	0.48-7.53	0.36	0		0.99

Tabella 10: Tabella incrociata della IU riportata dai clinici e dai pazienti a 3 mesi, e stratificati per età del paziente

Age		EHR a 3 mesi		Totale
		no IU	IU	
>=60 anni	N	10	9	19
	%	52.6%	47.4%	100.0%
<60 anni	N	8	28	36
	%	22.2%	77.8%	100.0%
Totale	N	18	37	55
	%	32.7%	67.3%	100.0%
>=60 anni	N	18	4	22
	%	81.8%	18.2%	100.0%
<60 anni	N	9	67	76
	%	11.8%	88.2%	100.0%
Totale	N	27	71	98
	%	27.6%	72.4%	100.0%

Discussione

Abbiamo sviluppato un flusso di estrazione e di data-mining che è in grado di catturare la documentazione fornita dai medici sugli esiti più importanti e centrati sul paziente nei pazienti con cancro alla prostata utilizzando note cliniche non strutturate e registrate di routine dalle EHR. Il nostro strumento ha identificato in modo efficiente e accurato le valutazioni dell'incontinenza urinaria dopo la prostatectomia e abbiamo riscontrato che la segnalazione da parte dei medici dei sintomi della IU è altamente concordante con quanto viene riferito dai pazienti utilizzando l'indagine EPIC-26, un importante PROM che nel corso degli anni è stato validato ed è attualmente considerato uno dei gold standard per la valutazione del paziente post-prostatectomia.

Questo studio dimostra quindi che i medici documentano accuratamente importanti PROMs nel testo narrativo non strutturato delle EHR e che questi risultati possono essere rilevati utilizzando uno strumento che sia in grado di fare data-mining dai documenti clinici a testo libero. Questo metodo automatizzato è un'importante conquista, che potrà avere un'ampia varietà di applicazioni in futuro e non si limita solo all'incontinenza urinaria né solamente ai risultati correlati al cancro. L'avanzamento della valutazione degli outcomes centrati sul paziente, insieme a tali metodologie automatizzate, potrà avere nel prossimo futuro effetti potenzialmente rilevanti sulla ricerca e la valutazione dei PROMs in generale.

I nostri risultati hanno diverse importanti implicazioni. Innanzitutto, abbiamo riscontrato un alto grado di concordanza tra i PROMs per incontinenza urinaria riportati in EPIC-26 con quanto registrato nella cartella clinica dopo estrazione con il nostro sistema di Natural Language Processing. In altri termini, questo indica come i risultati riferiti dal paziente ai loro medici (che vengono successivamente registrati nella cartella clinica) sembrano riflettere accuratamente l'esperienza dei pazienti su esiti che, come anche la letteratura conferma, sono molto più importanti e centrati sul paziente, sul

suo stato di salute e, non ultimo, sulla sua qualità di vita. È particolarmente degno di nota il fatto che questo alto grado di correlazione sia stato raggiunto nonostante la presenza di medici diversi che scrivevano le note cliniche, e quindi con diversi linguaggi e/o modi di dire o termini tecnici maggiormente usati.

I risultati dell'incontinenza urinaria riportati dall'indagine EPIC-26 sono simili e si associano bene, inoltre, con i risultati descrittivi di diversi studi recenti in cui IU è stata valutata nei pazienti, con EPIC-26, dopo prostatectomia radicale.(Barocas, 2017; Chen, 2017, Donovan, 2016) I nostri risultati sono rassicuranti, specialmente alla luce della preoccupazione che i pazienti sotto-riportino gli esiti avversi ai loro medici, come indicato da alcuni lavori recenti e che parlano di una “missing voice of the patients”. (Basch, 2010)

Esisteva un simile grado di concordanza tra dati di survey e dati riferiti dai pazienti per gli esiti urinari dopo l'intervento chirurgico in uno studio precedente. Tuttavia, per quanto è emerso ad oggi dalla letteratura nessuno studio ha ottenuto risultati di confronto e comparazione mediante estrazione diretta delle note cliniche degli EHR. (Wahlsh, 2000)

Il nostro successo nello sviluppo di uno strumento per l'estrazione automatizzata di esiti centrati sul paziente apre molte possibilità di applicazione ad altri tipi di dati. Il nostro strumento è altamente efficiente ed è in grado di valutare oltre 1 milione di note cliniche non strutturate in poche ore per specifiche domande sugli esiti di interesse. Tali applicazioni di data-mining possono essere applicate a diversi tipi di dati, siano essi strutturati e non strutturati. Ad esempio, dovrebbe essere possibile impostare sistemi automatici per l'estrazione di valori di laboratorio incorporati nelle note cliniche (ad es. valori rilevati dal medico attraverso esami che il paziente ha autonomamente condotto nel sistema privato), così come i livelli sierici di PSA. Questo permetterebbe di imputare moltissime informazioni sulle fasi cliniche o patologiche, e in questo modo tali dati possono essere trasferiti su

un database locale o eventualmente sui registri di patologia. Una volta che tali sistemi verranno sviluppati e convalidati, potrebbero fornire metodi più economici e potenzialmente più accurati per la raccolta e la unione di questi dati.

L'alto grado di correlazione tra i PROMs raccolti tramite strumento validato EPIC-26 e il metodo di data-mining che abbiamo sviluppato, apre opportunità per l'analisi dei PROMs da EHR anche molto ricche di informazioni e potenzialmente diverse come campo clinico, e non solo oncologico. Il nostro metodo consente l'incorporazione di codici diagnostici e procedurali per estrarre gli esiti, sebbene questi dati forniscano solo una componente delle informazioni critiche per una accurata valutazione degli esiti. Precedenti studi hanno dimostrato che la valutazione degli esiti utilizzando solo i codici diagnostici e di procedura per l'incontinenza urinaria non sono del tutto completi e spesso sono imprecisi. (Hernandez-Boussard, 2016)

Aggiungendo le informazioni critiche trovate nella parte narrativa, nei commenti scritti dai medici sulle cartelle cliniche, i livelli di precisione migliorano notevolmente. La nostra capacità di estrarre questi risultati apre la possibilità di valutare in modo efficiente e accurato qualsiasi PROMs, almeno per quanto riguarda l'incontinenza urinaria, in molti contesti di pratica clinica; questo può essere detto anche in considerazione del fatto che il nostro algoritmo è open source e può essere implementato in una varietà di EHR diverse. Ciò apre la possibilità di valutare i risultati centrati sul paziente sia a livello di pratica clinica, sia a livello di provider di servizi oppure, infine, a livello di popolazione.

Il nostro studio ha comunque alcuni limiti. In prima battuta, lo sviluppo di strumenti per la valutazione di specifici PROMs dipende da come sono prodotti i vocabolari per l'estrazione tramite NLP. In particolare si potrebbero riscontrare problematiche nell'adattamento di queste pratiche specialmente

nel trovare applicazione nei veri contesti pratici, nei quali i vocabolari potrebbero cambiare. Scopo di questo studio era anche quello di sviluppare questo fenotipo elettronico e questo algoritmo di estrazione per questo tipo di vocabolario al fine anche di una validazione esterna. Man mano che i sistemi sanitari utilizzeranno in futuro questi vocabolari, le modifiche dell'algoritmo potranno migliorare la precisione dell'estrazione e dare spazio a confronti fra diversi database e dati clinici. Va sottolineato come non abbiamo testato direttamente l'efficacia dell'estrazione dei dati della nostra NLP. Ad ogni modo, come detto nella sezione metodi, un infermiere del reparto di urologia ha revisionato manualmente 100 cartelle cliniche selezionate in modo random per sviluppare il vocabolario nel miglior modo possibile, il che ci dà fiducia sulla robustezza dell'estrazione della procedura impostata sull'estrazione tramite NLP. Secondariamente, il nostro lavoro dimostra come il nostro algoritmo funzioni bene per IU, potrebbe essere più difficile ottenere risultati più complessi come la disfunzione erettile o la qualità generale della vita. L'incontinenza urinaria può infatti essere quantificata (ad esempio: il numero di pannolini indossati al giorno) laddove altri esiti, come la disfunzione erettile, sono più soggettivi e difficili da misurare.

Infine, il nostro metodo si basa sulla documentazione da parte dei medici sugli esiti propri dei pazienti. In generale, nei sistemi sanitari occidentali, gli esiti riportati nel nostro sistema sanitario seguono bene quelli che vengono riportati anche dai pazienti; ma questo potrebbe non essere vero in altri contesti, con altre pratiche cliniche. Inoltre, i risultati riportati dai medici potrebbero essere influenzati dalla capacità di catturare direttamente dal testo libero della cartella clinica, in particolare se la sua compilazione è direttamente collegata alla struttura del rimborso. Tuttavia, dal momento che la documentazione dei medici sta diventando una componente di dati sempre più valida per fare analisi clinica, è probabile che in futuro sarà disponibile una documentazione testuale sempre più accurata e completa da

parte dei medici. Abbiamo deciso di comprimere le risposte utilizzando le istruzioni di EPIC-26 per calcolare il punteggio. Trasformare due variabili categoriali in dicotomiche potrebbe aver ridotto il potere della nostra analisi, ma è stato fatto poiché era l'unico modo per ottenere un confronto diretto tra le risposte a EPIC-26 e gli esiti valutati nelle EHR. Infine, abbiamo riscontrato una significativa perdita di pazienti che ci hanno restituito i PROMs al follow-up. Questa perdita significativa può aver limitato la forza dei nostri risultati.

In conclusione, comunque, abbiamo sviluppato uno strumento di data-mining per estrarre risultati di incontinenza urinaria dopo un intervento chirurgico per cancro alla prostata che mostra un alto grado di accordo con i risultati riportati dai pazienti utilizzando uno strumento validato. I metodi da me utilizzati per la formulazione di questa tesi sono tutti in grado di aprire importanti opportunità e scenari per lo sviluppo di database di dati, atti ad indirizzare la ricerca dei PROMs nei sistemi sanitari, con lo scopo ultimo di portare a processi di cura che siano sempre più adatti a rispondere alle esigenze dei pazienti.

Il lavoro futuro dei prossimi mesi si espanderà ad includere la valutazione di ulteriori parametri di qualità e risultati centrati sul paziente. In particolare i miei prossimi obiettivi di analisi includeranno l'analisi della disfunzione erettile, un esito di salute che impatta altrettanto, se non maggiormente, nella qualità di vita del paziente. Inoltre, è in programma una validazione delle performance degli algoritmi che abbiamo sviluppato in altre EHRs. Si sta programmando, attualmente, di effettuare un confronto diretto con l'InterMountain Healthcare System, una rete di ospedali privati statunitensi.

Conclusioni

Sulla base dei risultati della revisione della letteratura e della analisi comparativa, il dato che emerge più fortemente è che ancora c'è molto lavoro da fare per l'implementazione dei PROMs nei nostri contesti clinici.

Il dato della nostra revisione di letteratura sottolinea come, per la prostata in particolare (ma sicuramente il dato potrebbe essere trasferito in patologie di altri organi o altre branche al di fuori dell'oncologia) la misurazione della qualità sia sostanzialmente limitata alle misure di processo e di esito clinicamente *hard* quali mortalità e complicanze, più facilmente identificabili e misurabili, ma sicuramente meno impattanti sulla qualità della vita del paziente.

Risulta quindi fondamentale supportare in futuro la raccolta routinaria dei PROMs in vari contesti (anche al di fuori dell'oncologia) come metodo per mettere la "voce del paziente" al centro del processo di cura. La attuale mancanza di attenzione rispetto a queste misure deve condurre a miglioramenti sostanziali in termini di raccolta dati (*quali flussi informativi? Come integrati?*), confronto (*quale standard di riferimento, quali differenze significative nei confronti? In quali contesti assistenziali?*), utilizzazione (*per chi? in che modo?*), ed implementazione (*come li posso applicare in maniera sistematica ed utilizzarli per migliorare la qualità dell'assistenza erogata e realizzare modelli di cura patient-centered?*).

Ciò necessita la collaborazione multidisciplinare e trasversale dei diversi attori coinvolti nella attività di programmazione/gestione/erogazione delle cure, quindi, oltre che i pazienti, i clinici, i managers, gli amministratori/gestori dei servizi sanitari e i policy makers. Sarà quindi necessario un coinvolgimento di tutti gli stakeholder interessati e un loro maggiore coinvolgimento anche per la disseminazione di una cultura orientata al miglioramento della qualità con il paziente al centro.

In particolare, gli esperti del gruppo NHS per la valutazione dell'uso dei PROMs nel servizio sanitario inglese, sottolineano come le prospettive per la

implementazione e integrazione di questi indicatori nella valutazione della qualità dell'assistenza dovrà basarsi su:

- miglioramento dell'efficienza della raccolta dei dati, utilizzando l'acquisizione elettronica laddove appropriato;
- garanzia di chiarezza su come verranno utilizzati i dati. Questo deve essere esplicitato nelle comunicazioni e, in particolar modo, con i pazienti;
- fornitura di indicazioni ai providers sul modo migliore di interpretare e utilizzare i dati che saranno raccolti;
- valutazione del fatto che gli attuali incentivi finanziari basati sul rendimento possano rappresentare il metodo ottimale per incoraggiare il cambiamento del servizio;
- valutazione del monitoraggio da altre fonti con diversa granularità, e se, ad esempio le prestazioni a livello del professionista, possano valutare in modo più efficiente rispetto al modello esistente basato sul provider;
- verifica della presenza di una base solida di evidenze per i PROMs selezionati in ambito clinico, che siano definitivi e includano l'input del paziente.

Gli indicatori PROMs potranno essere utili utilizzati a livello *macro*, per monitorare le politiche sanitarie, a livello *meso* per migliorare la gestione dei servizi sanitari e a livello *micro* per migliorare la qualità dell'assistenza sui singoli pazienti. Un aspetto importante, può essere, a questo riguardo, il feedback in tempo reale dei dati PROMs per migliorare la comunicazione medico/paziente e anche informare le decisioni condivise.

Il caso di studio presentato (relativo alla valutazione della IU post intervento per carcinoma prostatico e alla sperimentazione e alla raccolta dati sugli esiti funzionali estratti dalle cartelle cliniche informatizzate utilizzando i dati del database Stanford University HealthCare), apre la strada alla possibilità di utilizzare e valutare questo tipo di indicatori in maniera sistematica traendo

informazioni dalla grande quantità di dati a disposizione in formato elettronico nell'ambito dei servizi sanitari. I clinici, infatti, stanti i nostri risultati nell'ambito degli esiti nella patologia oncologica prostatica (ma come da noi detto probabilmente questo dato sarà trasferibile anche in altri contesti) potrebbero riportare nelle cartelle cliniche importanti informazioni riferite dai pazienti. Ciò dovrà essere unito a programmi di formazione e di implementazione della raccolta dei PROMs in modo da convincere i clinici – che sono sicuramente più propensi ad affidarsi a *endpoints* “hard” (quali la morte, la sopravvivenza, etc.) – a valutare e a raccogliere nelle loro cartelle, in modo sistematico, informazioni riferite a questi esiti funzionali. Questo cambiamento potrà portare, in futuro, ad una migliore raccolta dei dati sugli esiti *patient-centered* e ad un effettivo miglioramento della qualità così orientata in tutti i settori della medicina.

Bibliografia:

- American Cancer Society. Cancer Facts & Figures 2016. Atlanta, Ga: American Cancer Society; 2016.
- Argimon JM, Limon E, Vila J, Cabezas C. Health-related quality of life in carers of patients with dementia. *Fam Pract.* 2004 Aug;21(4):454-7.
- Bakshi AB, Wee SL, Tay C, Wong LM, Leong IY, Merchant RA, Luo N. Validation of the care transition measure in multi-ethnic South-East Asia in Singapore. *BMC Health Serv Res.* 2012 Aug 16;12:256. doi: 10.1186/1472-6963-12-256.
- Barocas DA, Alvarez J, Resnick MJ, et al. Association Between Radiation Therapy, Surgery, or Observation for Localized Prostate Cancer and Patient-Reported Outcomes After 3 Years. *JAMA.* 2017 Mar 21;317(11):1126-1140.
- Basch E. The missing voice of patients in drug-safety reporting. *N Engl J Med.* 2010;362(10):865-
- Bausewein C, Fegg M, Radbruch L, Nauck F, von Mackensen S, Borasio GD, Higginson IJ. Validation and clinical application of the german version of the palliative care outcome scale. *J Pain Symptom Manage.* 2005 Jul;30(1):51-62.
- Bekelman JE, Zelefsky MJ, Jang TL, Basch EM, Schrag D. Variation in adherence to external beam radiotherapy quality measures among elderly men with localized prostate cancer. *Int J Radiat Oncol Biol Phys.* 2007 Dec 1;69(5):1456-66.
- Blumenthal D, Malphrus E, McGinnis JM, editors. *Vital signs: core metrics for health and health care progress.* Washington (DC): National Academies Press; 2015. p. B-9.
- Bosworth A, Cox M, O'Brien A, Jones P, Sargeant I, Elliott A, Bukhari M. Development and Validation of a Patient Reported

Experience Measure (PREM) for Patients with Rheumatoid Arthritis (RA) and other Rheumatic Conditions. *Curr Rheumatol Rev.* 2015;11(1):1-7.

- Boyce MB, Browne JP. Does providing feedback on patient-reported outcomes to healthcare professionals result in better outcomes for patients? A systematic review. *Qual Life Res.* 2013 Nov;22(9):2265-78.
- Boyce MB, Browne JP, Greenhalgh J. Surgeon's experiences of receiving peer benchmarked feedback using patient-reported outcome measures: a qualitative study. *Implement Sci.* 2014 Jun 27;9:84.
- Boyce MB, Browne JP. The effectiveness of providing peer benchmarked feedback to hip replacement surgeons based on patient-reported outcome measures--results from the PROFILE (Patient-Reported Outcomes: Feedback Interpretation and Learning Experiment) trial: a cluster randomised controlled study. *BMJ Open.* 2015 Jul 31;5(7):e008325.
- Boyer L, Lançon C, Baumstarck K, Parola N, Berbis J, Auquier P. Evaluating the impact of a quality of life assessment with feedback to clinicians in patients with schizophrenia: randomised controlled trial. *Br J Psychiatry.* 2013 Jun;202:447-53.
- Brundage M, Danielson B, Pearcey R, Bass B, Pickles T, Bahary JP, Peng Y, Wallace D, Mackillop W. A criterion-based audit of the technical quality of external beam radiotherapy for prostate cancer. *Radiother Oncol.* 2013 Jun;107(3):339-45. doi: 10.1016/j.radonc.2013.04.023. Epub 2013 Jul 2.
- Carpenter I, Hirdes JP Using InterRAI Assessment Systems to Measure and Maintain Quality of Long-term Care , *A Good Life in Old Age? Monitoring and Improving Quality in Long-term Care*, 2013 OECD/European Union, OECD Publishing, Paris,

- Casalino LP, Gans D, Weber R, Cea M, Tuchovsky A, Bishop TF, Miranda Y, Frankel BA, Ziebler KB, Wong MM, Evenson TB. US physician practices spend more than \$15.4 billion annually to report quality measures. *Health Affairs*, 2016 35(3):401-406.
- Chen RC, Basak R, Meyer AM, et al. Association Between Choice of Radical Prostatectomy, External Beam Radiotherapy, Brachytherapy, or Active Surveillance and Patient-Reported Quality of Life Among Men With Localized Prostate Cancer. *JAMA*. 2017 Mar 21;317(11):1141-1150.
- Chen J, Ou L, Hollis SJ. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Serv Res*. 2013 Jun 11;13:211.
- Chien AT, Rosenthal MB Medicare's physician value-based payment modifier—will the tectonic shift create waves? *New England Journal of Medicine*, 2013 369(22), 2076-2078.
- CIHI – Canadian Institute for Health Information (2015), “PROMs Background Document”.
- Coleman EA, Mahoney E, Parry C. Assessing the quality of preparation for posthospital care from the patient's perspective: the care transitions measure. *Med Care*. 2005 Mar;43(3):246-55.
- Cooperberg MR, Vickers AJ, Broering JM, et al: Comparative risk-adjusted mortality outcomes after primary surgery, radiotherapy, or androgen-deprivation therapy for localized prostate cancer. *Cancer* 116:5226-5234, 2010
- A. Cooperberg MR, Broering JM, Carroll PR: Time trends and local variation in primary treatment of localized prostate cancer. *J Clin Oncol* 28:1117-1123, 2010

- Devlin, N.J. Getting the Most out of PROMs: Putting Health Outcomes at the Heart of NHS Decision-Making, 2010 The King's Fund.
- D'Hoore W, Sicotte C, Tilquin C. Risk adjustment in outcome assessment: the Charlson comorbidity index. *Methods Inf Med* 1993;32(5):382-7.
- Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q.* 1966 Jul;44(3):Suppl:166-206.
- Donabedian A. The Quality of care. How can it be assessed? *JAMA* 1988; 260: 1743-8
- Donabedian A. The seven pillars of quality. *Arch Pathol Lab Med.* 1990 Nov;114(11):1115-8. Fremont AM, Cleary PD, Hargraves JL, Rowe RM, Jacobson NB, Ayanian JZ. Patient-centered processes of care and long-term outcomes of myocardial infarction. *J Gen Intern Med.* 2001 Dec;16(12):800-8.
- Donovan JL, Hamdy FC, Lane JA, et al. Patient-Reported Outcomes after Monitoring, Surgery, or Radiotherapy for Prostate Cancer. *N Engl J Med.* 2016 Oct 13;375(15):1425-1437.
- Eggener SE, Scardino PT, Walsh PC, et al: Predicting 15-year prostate cancer specific mortality after radical prostatectomy. *J Urol* 185:869-875, 2011
- Etzioni R, Gulati R, Tsodikov A, et al. The prostate cancer conundrum revisited. *Cancer* 2012; 118:5955 – 5963
- Flynn BS, Wood ME, Ashikaga T, Stockdale A, Dana GS, Naud S. Primary care physicians' use of family history for cancer risk assessment. *BMC Fam Pract.* 2010 Jun 3;11:45.
- Friedberg MW, SteelFisher GK, Karp M, Schneider EC. Physician groups' use of data from patient experience surveys. *J Gen Intern Med.* 2011 May;26(5):498-504.

- D'Avolio LW, Litwin MS, Rogers SO Jr, Bui AA. Facilitating Clinical Outcomes Assessment through the automated identification of quality measures for prostate cancer surgery. *J Am Med Inform Assoc.* 2008 May-Jun;15(3):341-8. doi: 10.1197/jamia.M2649.
- Flores LT, Bennett AV, Law EB, Hajj C, Griffith MP, Goodman KA. Patient-Reported Outcomes vs. Clinician Symptom Reporting During Chemoradiation for Rectal Cancer. *Gastrointest Cancer Res* 2012;5(4):119-24.
- Fujisawa R, Klazinga N. "Measuring Patient Experiences (PREMs): Progress Made by the OECD and its Member Countries 2006-2015", 2017 – In press, OECD Publishing, Paris.
- Garratt AM, Solheim E, Danielsen K "National and Crossnational Surveys of Patient Experiences: A Structured Rreview", Rapport No. 7-2008, Nasjonalt kunnskapssenter for helsetjenesten (Norwegian Knowledge Centre for the Health Services), Oslo.
- Ghani KR, Miller DC. Variation in prostate cancer care. *JAMA*, 2015 313(20):2066-7.
- Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ.* 2005;331(7524):1064-1065
- Greenhalgh J, Pawson R, Wright J, Black N, Valderas JM, Meads D, Gibbons E, Wood L, Wood C, Mills C, Dalkin S Functionality and feedback: a protocol for a realist synthesis of the collation, interpretation and utilisation of PROMs data to improve patient care. *BMJ Open.* 2014 Jul 22;4(7):e005601.
- Harkema H, Dowling JN, Thornblade T, Chapman WW. ConText: an algorithm for determining negation, experiencer, and temporal status from clinical reports. *J Biomed Inform* 2009;42(5):839-51.
- Harlan L, Brawley O, Pommerenke F, Wali P, Kramer B. Geographic, age, and racial variation in the treatment of local/regional

carcinoma of the prostate. *Journal of Clinical Oncology*, 1995 13(1):93-100.

- Haverman L, van Rossum MA, van Veenendaal M, van den Berg JM, Dolman KM, Swart J, Kuijpers TW, Grootenhuis MA. Effectiveness of a web-based application to monitor health-related quality of life. *Pediatrics*. 2013 Feb;131(2):e533-43.
- Heffner JE, Holgate ST, Chung KF, Niederman MS, Daley CL, Jett JR, et al. Road ahead to respiratory health: experts chart future research directions. *Respirology* 2009;14(5):625-36.
- Hearn J, Higginson IJ. Development and validation of a core outcome measure for palliative care: the palliative care outcome scale. Palliative Care Core Audit Project Advisory Group. *Qual Health Care*. 1999 Dec;8(4):219-27.
- Hernandez-Boussard T, Tamang S, Blayney D, Brooks J, Shah N. New Paradigms for Patient-Centered Outcomes Research in Electronic Medical Records: An Example of Detecting Urinary Incontinence Following Prostatectomy. *EGEMS (Wash DC)*. 2016 May 12;4(3):1231.
- Hernandez-Boussard T, Kourdis P, Seto T, Ferrari M, Blayney DW, Rubin D, et al., editors. Mining Electronic Health Records to Extract Patient-Centered Outcomes Following Prostate Cancer Treatment. *AMIA Annual Symposium proceedings / AMIA Symposium*; 2017 Nov 2, 2017.
- Herrel LA, Kaufman SR, Yan P, Miller DC, Schroeck FR, Skolarus TA, Shahinian VB, Hollenbeck BK. Health Care Integration and Quality among Men with Prostate Cancer. *J Urol*. 2017 Jan;197(1):55-60.
- Holmes JA, Bensen JT, Mohler JL, Song L, Mishel MH, Chen RC. Quality of care received and patient-reported regret in prostate

cancer: Analysis of a population-based prospective cohort. *Cancer*. 2017 Jan 1;123(1):138-143.

- Howell D, Molloy S, Wilkinson K, Green E, Orchard K, Wang K, Liberty J. Patient-reported outcomes in routine cancer clinical practice: a scoping review of use, impact on health outcomes, and implementation factors. *Ann Oncol*. 2015 Sep;26(9):1846-58.
- Jenkinson C, Fitzpatrick R, Peto V, Greenhall R, Hyman N. The Parkinson's Disease Questionnaire (PDQ-39): development and validation of a Parkinson's disease summary index score. *Age Ageing*. 1997 Sep;26(5):353-7.
- Jha, AK Time to get serious about pay for performance. *JAMA*, 2013, 309(4), 347-348.
- Jha AK, Pronovost P Toward a safer health care system: the critical need to improve measurement. *JAMA*, 2016 3; 315 (17):1831-2.
- Kahn KL, Schneider EC, Malin JL, Adams JL, Epstein AM. Patient centered experiences in breast cancer: predicting long-term adherence to tamoxifen use. *Med Care*. 2007 May;45(5):431-9.
- Kamal AH, Bull J, Stinson CS, Blue DL, Abernethy AP. Conformance with supportive care quality measures is associated with better quality of life in patients with cancer receiving palliative care. *J Oncol Pract*. 2013 May;9(3):e73-6.
- Kowalczyk KJ, Levy JM, Caplan CF, Lipsitz SR, Yu HY, Gu X, et al. Temporal national trends of minimally invasive and retropubic radical prostatectomy outcomes from 2003 to 2007: results from the 100% Medicare sample. *Eur Urol* 2012;61(4):803-9.
- ICHOM, "The Right Thing for Patients: Pioneering Outcomes Measurement in Wales", 2015 Cambridge, United States, available at: <http://www.ichom.org/others/the-right-thing-for-patients-pioneeringoutcomes-measurement-in-wales/> Institute of Medicine.

Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academies Press; 2001.

- InterRAI, "Quality of Life (QoL)", 2015; available at: <http://interrai.org/quality-of-life.html>
- Kirby JC, Speltz P, Rasmussen LV, Basford M, Gottesman O, Peissig PL, et al. PheKB: a catalog and workflow for creating electronic phenotype algorithms for transportability. *J Am Med Inform Assoc* 2016;23(6):1046-52.
- Kyte D, Ives J, Draper H, Calvert M. Current practices in patient-reported outcome (PRO) data collection in clinical trials: a cross-sectional survey of UK trial staff and management. *BMJ Open*. 2016 Oct 3;6(10):e012281.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159-74.
- Lee DJ, Consedine NS, Gonzalez JR, Spencer BA. Association of healthcare barriers with prostate-specific antigen screening among African-American and Afro-Caribbean men. *Urology*. 2012 Sep;80(3):556-63.
- Litwin MS, Steinberg M, Malin JL, et al. 2000; Prostate cancer patient outcomes and choice of providers: Development of an infrastructure for quality assessment. Santa Monica (CA): RAND
- Litwin MS, Hays RD, Fink A, et al. Quality-of-life outcomes in men treated for localized prostate cancer. *JAMA: the journal of the American Medical Association*. 1995;273(2):129–135.
- Lohr KN. Medicare: a strategy for quality assurance. *J Qual Assur*. 1991 Jan-Feb;13(1):10-3.
- Makarov DV, Desai R, Yu JB, Sharma R, Abraham N, Albertsen PC, Krumholz HM, Penson DF, Gross CP. Appropriate and inappropriate imaging rates for prostate cancer go hand in hand by

region, as if set by thermostat. *Health Aff (Millwood)*. 2012 Apr;31(4):730-40.

- Miller DC, Sanda MG, Dunn RL, et al. Long-term outcomes among localized prostate cancer survivors: health-related quality-of-life changes after radical prostatectomy, external radiation, and brachytherapy. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*. 2005;23(12):2772–2780.
- Miller DC, Spencer BA, Ritchey J, Stewart AK, Dunn RL, Sandler HM, Wei JT, Litwin MS. Treatment choice and quality of care for men with localized prostate cancer. *Med Care*. 2007 May;45(5):401-9.
- Monitor & NHS England. 2014/15 National Tariff Payment System. London: Monitor & NHS England, 2013.
- Musen MA, Noy NF, Shah NH, Whetzel PL, Chute CG, Story MA, et al. The National Center for Biomedical Ontology. *J Am Med Inform Assoc* 2012;19(2):190-
- Nowels D, Kamerow, DB New “Core Quality Measures”: Only a Beginning. *The Journal of the American Board of Family Medicine*, 2017 30(1), 4-7.
- O'Connor NR, Hu R, Harris PS, Ache K, Casarett DJ. Hospice admissions for cancer in the final days of life: independent predictors and implications for quality measures. *J Clin Oncol*. 2014 Oct 1;32(28):3184-9.
- Panzer RJ, Gitomer RS, Greene WH , Webster PR, Landry KR, Riccobono CA Increasing demands for quality measurement. *JAMA*, 2013, 310(18), 1971-1980.
- Parry C, Mahoney E, Chalmers SA, Coleman EA. Assessing the quality of transitional care: further applications of the care transitions measure. *Med Care*. 2008 Mar;46(3):317-22.
- Partridge T, Carluke I, Emmerson K, Partington P, Reed M. Improving patient reported outcome measures (PROMs) in total knee

replacement by changing implant and preserving the infrapatella fatpad: a quality improvement project. *BMJ Qual Improv Rep.* 2016 May 5;5(1).

- Penson DF, McLerran D, Feng Z, et al. 5-year urinary and sexual outcomes after radical prostatectomy: results from the prostate cancer outcomes study. *The Journal of urology.* 2005;173(5):1701–1705.
- Peters M, Jenkinson C, Doll H, Playford ED, Fitzpatrick R. Carer quality of life and experiences of health services: a cross-sectional survey across three neurological conditions. *Health Qual Life Outcomes.* 2013 Jun 25;11:103.
- Ranpura V, Agrawal S, Chokshi P, Yerasi C, Wood L, Beasley J, Harkness J, Heller S, Self L, Swain SM, Jang S. Improving Documentation of Pain Management at MedStar Washington Cancer Institute. *Oncol Pract.* 2015 Mar;11(2):155-7.
- Report AHRQ, “aligning forces for quality”, Settembre 2010 Disponibile liberamente al sito: <https://www.issuelab.org/resources/6149/6149.pdf>
- Report NQF: Improving Diagnostic quality and safety: Draft Measurement Framework. National Quality Forum. Jan 2017.
- Report OECD, “Recommendations from high level reflection group on the future of health statistics”, Gennaio 2017 Disponibile liberamente al sito: <https://www.oecd.org/els/health-systems/Recommendations-from-high-level-reflection-group-on-the-future-of-health-statistics.pdf>
- NHS England. National Patient Reported Outcome Measures (PROMs) Programme Consultation. Disponibile a: <https://www.engage.england.nhs.uk/consultation/proms-programme>.
- Samuel CA, Landrum MB, McNeil BJ, Bozeman SR, Williams CD, Keating NL. Racial disparities in cancer care in the Veterans Affairs

health care system and the role of site of care. *Am J Public Health*. 2014 Sep;104 Suppl 4:S562-71.

- Sanda MG, Dunn RL, Michalski J, et al. Quality of life and satisfaction with outcome among prostate-cancer survivors. *N Engl J Med*. 2005;358(12):1250–1261.
- Schroeck FR, Kaufman SR, Jacobs BL, Hollenbeck BK. Receipt of best care according to current quality of care measures and outcomes in men with prostate cancer. *J Urol*. 2015 Feb;193(2):500-4.
- A. Schroeck FR, Kaufman SR, Jacobs BL, Skolarus TA, Miller DC, Montgomery JS, Weizer AZ, Hollenbeck BK. Adherence to performance measures and outcomes among men treated for prostate cancer. *J Urol*. 2014 Sep;192(3):743-8. doi: 10.1016/j.juro.2014.03.091.
- B. Schroeck FR, Kaufman SR, Jacobs BL, Skolarus TA, Hollingsworth JM, Shahinian VB, Hollenbeck BK Regional variation in quality of prostate cancer care. *J Urol*. 2014 Apr;191(4):957-62.
- C. Schroeck FR, Kaufman SR, Jacobs BL, Skolarus TA, Zhang Y, Hollenbeck BK. Technology diffusion and prostate cancer quality of care. *Urology*. 2014 Nov;84(5):1066-72.
- Sequist TD, Schneider EC, Anastario M, Odigie EG, Marshall R, Rogers WH, Safran DG. Quality monitoring of physicians: linking patients' experiences of care to clinical quality and outcomes. *J Gen Intern Med*. 2008 Nov;23(11):1784-90. doi: 10.1007/s11606-008-0760-4. Epub 2008 Aug 28.
- Shekelle P, Lim Y-W, Mattke S, Damberg C Does public release of performance results improve quality of care. A systematic review. London: The Health Foundation, 2008.
- Shelton JB, Skolarus TA, Ordin D, Malin J, Antonio A, Ryoo J, Saigal CS. Validating electronic cancer quality measures at Veterans Health Administration. *Am J Manag Care*. 2014;20(12):1041-7.

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016; 66:7 – 30
- Skolarus TA, Chan S, Shelton JB, Antonio AL, Sales AE, Malin JL, Saigal CS. Quality of prostate cancer care among rural men in the Veterans Health Administration. *Cancer*. 2013 Oct 15;119(20):3629-35.
- Smith SC, Lamping DL, Banerjee S, Harwood R, Foley B, Smith P, Cook JC, Murray J, Prince M, Levin E, Mann A, Knapp M. Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology. *Health Technol Assess*. 2005 Mar;9(10):1-93, iii-iv.
- Snowden, A. (2015), “Health Care Quality Report Minneapolis”, Minnesota Community Measurement.
- Sohn W, Resnick MJ, Greenfield S, Kaplan SH, Phillips S, Koyama T, Goodman M, Hamilton AS, Hashibe M, Hoffman KE, Paddock LE, Stroup AM, Wu XC, Penson DF, Barocas DA. Impact of Adherence to Quality Measures for Localized Prostate Cancer on Patient-reported Health-related Quality of Life Outcomes, Patient Satisfaction, and Treatment-related Complications. *Med Care*. 2016 Aug;54(8):738-44.
- Stanford JL, Feng Z, Hamilton AS, et al. Urinary and sexual function after radical prostatectomy for clinically localized prostate cancer: the Prostate Cancer Outcomes Study.
- Strömquist B, Fritzell P, Hägg O, Jönsson B, Sandén B; Swedish Society of Spinal Surgeons. Swespine: the Swedish spine register : the 2012 report. *Eur Spine J*. 2013 Apr;22(4):953-74.
- Szymanski KM, Wei JT, Dunn RL, Sanda MG. Development and validation of an abbreviated version of the expanded prostate cancer index composite instrument for measuring health-related quality of life among prostate cancer survivors. *Urology*. 2010;76(5):1245-50.

- Takeuchi EE, Keding A, Awad N, Hofmann U, Campbell LJ, Selby PJ, Brown JM, Velikova G. Impact of patient-reported outcomes in oncology: a longitudinal analysis of patient-physician communication. *J Clin Oncol.* 2011 Jul 20;29(21):2910-7.
- Toole M, Lutz S, Johnstone PA. Radiation oncology quality: aggressiveness of cancer care near the end of life. *J Am Coll Radiol.* 2012 Mar;9(3):199-202.
- Van Noorden M Routine Outcome Monitoring in Psychiatric Clinical Practice: Background, Overview and Implications for Personcentered Psychiatry *Eur J Pers Cent Healthcare*, 2013, 1(1):103-111
- Varagunam M, Hutchings A, Neuburger J and Black N. Impact on hospital performance of introducing routine patient reported outcome measures in surgery. *J Health Serv Res Policy* 2014; 19: 77–84.
- Walsh PC, Marschke P, Ricker D, Burnett AL. Patient-reported urinary continence and sexual function after anatomic radical prostatectomy. *Urology* 2000;55(1):58-61.
- Webster TM, Newell C, Amrhein JF, Newell KJ. Cancer Care Ontario Guidelines for radical prostatectomy: striving for continuous quality improvement in community practice. *Can Urol Assoc J.* 2012 Dec;6(6):442-5. doi: 10.5489/cuaj.10195.
- A. Wei JT, Dunn RL, Sandler HM, McLaughlin PW, Montie JE, Litwin MS, Nyquist L, Sanda MG. Comprehensive comparison of health-related quality of life after contemporary therapies for localized prostate cancer. *J Clin Oncol.* 2002 Jan 15;20(2):557-66.
- B. Wei JT. Scoring Instructions for the Expanded Prostate cancer Index Composite (EPIC). *Ann Arbor* 2002;1001:48109-0330.
- Wei JT, Dunn RL, Litwin MS, Sandler HM, Sanda MG. Development and validation of the expanded prostate cancer index composite (EPIC) for comprehensive assessment of health-related quality of life in men with prostate cancer. *Urology* 2000;56(6):899-905.

- Worwag E, Chodak GW. Overnight hospitalization after radical prostatectomy: the impact of two clinical pathways on patient satisfaction, length of hospitalization, and morbidity. *Anesth Analg.* 1998 Jul;87(1):62-7.
- Whiting PF, Moore TH, Jameson CM, et al. Symptomatic and quality-of-life outcomes after treatment for clinically localized prostate cancer: a systematic review. *BJU Int.*2016;118(2):193-204.
- Wright JD, Tergas AI, Ananth CV, Burke WM, Chen L, Neugut AI, Richards CA, Hershman DL Relationship between surgical oncologic outcomes and publically reported hospital quality and satisfaction measures.. *J Natl Cancer Inst.* 2015 Jan 24;107(3).
- Xian Y, Wu J, O'Brien EC, Fonarow GC, Olson DM, Schwamm LH, Bhatt DL, Smith EE, Suter RE, Hannah D, Lindholm B, Maisch L, Greiner MA, Lytle BL, Pencina MJ, Peterson ED, Hernandez AF. Real world effectiveness of warfarin among ischemic stroke patients with atrial fibrillation: observational analysis from Patient-Centered Research into Outcomes Stroke Patients Prefer and Effectiveness Research (PROSPER) study. *BMJ.* 2015 Jul 31;351:h3786.