

VIII Rapporto *Sanità - CEIS*

*Opzioni di Welfare
e integrazione delle politiche*



Università degli Studi
di Roma "Tor Vergata"

5b - Indicatori di performance: un indice unico di performance dei SSR

Daniela d'Angela¹, Federico Spandonaro¹

5b.1. Introduzione

Negli ultimi anni il processo di valutazione dei risultati o *performance* ha avuto una forte accelerazione, spinta da una apprezzabile volontà, in particolare, del Ministero della Salute, di dotarsi e dotare i Servizi Sanitari Regionali (SSR) di strumenti di *accountability*.

Si può serenamente affermare che già oggi il settore sanitario è largamente avanti rispetto alle altre aree della Pubblica Amministrazione nei processi valutativi.

Tanto per citare due esperienze particolarmente significative, il Ministero supporta la conoscenza dello stato del sistema sanitario con il sistema di monitoraggio sviluppato dalla Scuola Superiore S. Anna di Pisa, noto come i "bersagli", ed il recente Piano Nazionale di Valutazione Esiti (PNE) gestito da Age.Na.S.

Non si può che esprimere il proprio plauso per questo, invero nuovo, corso del SSN, ed augurarsi che la possibilità di valutazione si estenda alle molte aree, *in primis* quella cosiddetta territoriale, che ne sono in pratica ancora sprovviste.

Allo stesso tempo la crescita (anche numerica) dei modelli di valutazione, caratterizzati da criteri e obiettivi diversi, se rende conto della complessa multi-dimensionalità del comparto sanitario, comporta il rischio di una perdita di sintesi. Si pone, in altri termini, il problema di utilizzare le crescenti informazioni disponibili, per addivenire ad una valutazione complessiva delle *performance* dei sistemi regionali, specie ora che, affermandosi il federalismo, si moltiplicano le diversità nell'approccio alla tutela del diritto alla salute.

Fatta salva l'importanza da attribuire all'opportunità derivante dal fatto che ogni singolo metodo offre agli amministratori dei sistemi regionali (e/o aziendali) informazioni utili per migliorare il proprio operato, appare utile poter esprimere valutazioni di sintesi: appare, infatti evidente, che ogni sistema può, in via di principio eccellere su un aspetto e non su un altro; né si può negare che in qualche misura si possano creare *trade-off* fra obiettivi: non a caso, a titolo di mero esempio, a livello internazionale si avverte una forte regolarità nella relazione (inversa) fra spesa e *responsiveness* dei sistemi sanitari.

La sintesi è resa peraltro complessa dal fatto che è presumibile (e, probabilmente, ragionevolmente auspicabile) che *stakeholder* diversi possano attribuire (a titolo personale, ma ancor di più in quanto portatori di interessi di categoria diversi), scale di preferenza diverse ai vari obiettivi che si possono configurare.

¹ CEIS Sanità, Facoltà di Economia, Università degli Studi di Roma Tor Vergata

Per questo si ritiene che una sintesi valutativa in campo sanitario non possa basarsi su una esplicita, o peggio implicita, applicazione di criteri tecnocratici, bensì debba determinarsi sulla base di un metodo partecipativo e democratico.

Ci sembra di essere suffragati in questa proposizione, dall'impostazione stessa della legge istitutiva del SSN, come anche dall'afflato posto dalle riforme degli anni '90 sul ruolo del cittadino/utente, ovvero sul principio di sovranità del consumatore.

In quanto segue, quindi, si esplicita una proposta metodologica per giungere alla costruzione di un processo valutativo dei SSR (replicabile a livello intraregionale per valutare le *performance* aziendali), che abbia caratteristiche di semplicità, immediatezza e, prioritariamente, partecipazione democratica.

5b.2 La valutazione della *performance*: lineamenti generali

I due fattori più critici nell'individuazione di un indice unico di *performance* sono: la scelta di indicatori che siano rappresentativi della *performance* dei Sistemi Sanitari e, quindi, la scelta della funzione di composizione, ovvero l'importanza relativa che si attribuisce loro.

Alla questione della scelta degli indicatori, è legata la questione della loro standardizzazione, operazione assolutamente imprescindibile affinché gli indicatori siano confrontabili e quindi atti a effettuare operazioni di *benchmarking*; la standardizzazione va intesa come metodo per depurare gli indicatori da ogni elemento di variabilità dovuta a determinanti non considerate rilevanti ai fini della valutazione della *performance*: nel nostro caso tipicamente, ma non esclusivamente, le differenze di bisogno sanitario delle diverse popolazioni regionali (per esemplificare, una spesa maggiore può non significare minore efficienza, se è stata sostenuta a fronte di una popolazione più anziana; analogamente, un esito peggiore può dipendere da una maggiore presenza di fattori di rischio, etc.).

Non è neppure da trascurare l'aspetto delle caratteristiche e proprietà degli indicatori: criteri quali la non ridondanza, la specificità etc. sono, seppure con diverse intensità e cogenze, richiesti praticamente da qualsiasi tecnica di composizione, e quindi andrebbero esplicitati e analizzati.

Una cosa che ci sembra importante sottolineare è la non ridondanza: utilizzare un numero elevato di indicatori non aggiunge precisione al modello di decisione; si rischia piuttosto una ridondanza che distorce i risultati in favore delle dimensioni per cui sono stati previsti più indicatori. Passando all'aspetto della composizione, e senza entrare in un'analisi approfondita del concetto di *performance* di un Sistema Sanitario, ci limitiamo qui a ricordare come la letteratura sia sufficientemente concorde nel ritenere che la *performance* abbia una natura multi-dimensionale, in ossequio alla complessità unanimemente riconosciuta al settore sanitario.

Da questa assunzione derivano, quindi, almeno due conseguenze fra loro interrelate.

La prima, relativa all'identificazione delle diverse dimensioni della *performance*, in modo da garantire che esse siano ragionevolmente rappresentate dagli indicatori.

Per *performance* sostanzialmente si intende una sintesi, una sorta di media (composizione) dei singoli risultati ottenuti in ogni singola dimensione: ad es. efficienza, efficacia e qualità possono essere diverse dimensioni della *performance* complessiva.

La seconda conseguenza è che ogni dimensione è, in generale, misurata in unità naturali diverse, in via di principio non direttamente "sommabili" fra loro; quindi per addivenire ad una quan-

tificazione della *performance* complessiva, è necessario ricondurle ad un'unica unità di misura, indipendente dalle caratteristiche di ogni singola dimensione.

Un ulteriore elemento di riflessione è dato dalla natura dei sistemi sanitari oggetto di analisi: nel caso specifico italiano, la natura pubblica dei SSR va tenuta in conto, condizionando tanto le dimensioni (ad esempio l'equità è una dimensione propria della *performance* di un servizio pubblico e non necessariamente lo sarebbe di uno privatistico), quanto la prospettiva dell'analisi.

Nello scenario internazionale si annoverano diverse esperienze di definizione di un indice unico di *performance* in ambito sanitario, tra cui citiamo quella statunitense, canadese, britannica, quella della *World Health Organization* (WHO). La metodologia introdotta dall'*United States Medicare* nel 2000 e 2003 si basa sulla composizione di ventidue indicatori di pari peso distribuiti in sei aree cliniche; quella del *Canadian Regional Health Care* si basa sulla composizione di quindici indicatori distribuiti in sei categorie (*Outcomes, Prenatal care, Community health, Elderly services Efficiency e Resources*): il peso degli indicatori all'interno della categoria è stato definito attraverso un sistema di *expert-opinion*, la composizione delle sei categorie è invece ottenuta con un sistema di pesi definito in modo arbitrario. Il Sistema britannico nel 2000 ha invece sviluppato un indice unico di *performance* basato sulla composizione di sei indicatori utilizzando un sistema di pesi basato sulle preferenze pubbliche, rilevate con una *survey* applicando il metodo "budget-pie", ovvero chiedendo l'allocazione di 60 *chips* sui sei indicatori (Appleby & Mulligan, 2000). Ancora citiamo quello implementato nel 2000 dalla WHO su 191 Paesi (*World Health Report*, 2000), la cui composizione si basa su una somma dei *ranking* di indicatori afferenti a quattro macroaree (*Overall health outcomes, Inequality in health, Fairness of financing e Health system responsiveness*). I pesi sono stati definiti attraverso una *survey* effettuata su 2000 figure "chiave", cui veniva richiesto di attribuire un punteggio ad ogni singolo indicatore.

5b.3 L'approccio basato sulla utilità prodotta

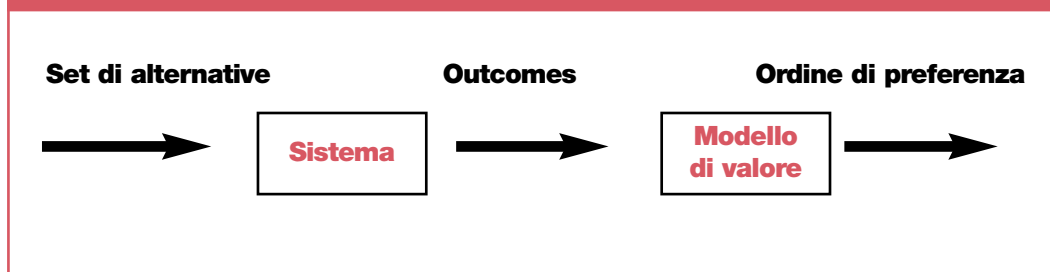
La metodologia che viene qui proposta, intende replicare un processo decisionale, così come avviene nella teoria della domanda in Economia; si fonda, quindi, sull'idea che gli *stakeholder* del sistema perseguano diversi obiettivi (dimensioni), a cui attribuiscono in base alle loro personali preferenze, un grado di utilità differente.

Gli *stakeholder* definiscono altresì un saggio marginale di sostituzione degli obiettivi, ovvero ritengono che una migliore *performance* su una dimensione almeno parzialmente possa compensare una peggiore *performance* su un altro.

In pratica si assume che non tutti gli obiettivi forniscono la stessa utilità, e che possono essere almeno parzialmente sostituiti uno con l'altro nella composizione della *performance* complessiva.

La tecnica adottata elicitava una funzione di utilità dalle preferenze dei singoli *stakeholder* (e poi come vedremo genera la funzione di utilità sociale), da cui risulta possibile stimare il livello relativo di utilità fornita dalla *performance* di ogni singolo SSR, esplicitando un *ranking* corrispondente alla struttura delle preferenze sottostanti.

Figura 5b.1 - Paradigma delle analisi decisionali



Come premesso, obiettivo della presente proposta è quello di addivenire ad un processo decisionale di tipo democratico e partecipativo. Ovviamente, trattandosi di servizi pubblici, al centro del sistema avremo le preferenze dei cittadini (rappresentati dalle loro associazioni di *advocacy* e tutela, e dai politici); argomentazioni relative ai noti problemi di asimmetria informativa che caratterizzano il settore sanitario ci sembra che ragionevolmente giustifichino la considerazione nel processo decisionale delle preferenze “informate” di tecnici, quali i professionisti sanitari (clinici, infermieri, etc.) e i manager dei sistemi sanitari (nazionali, ivi compreso il Ministero della Salute, regionali e aziendali); il notevole impatto economico dei sistemi sanitari giustifica, a parere di chi scrive, anche la considerazione nel processo decisionale delle preferenze dei rappresentanti del Ministero dell’Economia, dei *provider* e dei rappresentanti delle industrie afferenti al settore sanitario (dispositivi medici, farmaceutica, etc.).

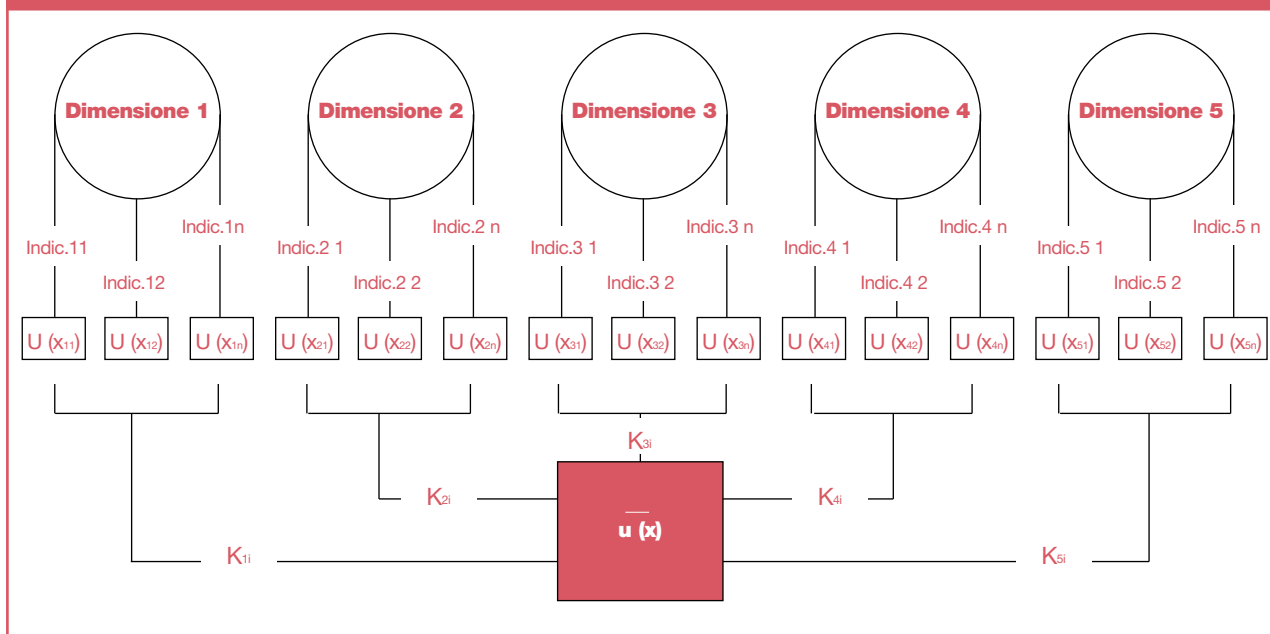
Un pregio del metodo proposto è anche quello di permettere un’elicitazione di preferenze per sottogruppo, rendendo possibile evidenziare diversi *ranking* e quindi confrontare le diverse valutazioni della *performance* espresse da gruppi diversi di *stakeholder*.

Si pone, piuttosto, il problema di addivenire ad una aggregazione delle preferenze che generi un sistema di “preferenze sociali”.

In pratica, nell’implementazione del metodo viene normalmente utilizzato il metodo della media aritmetica (*Harsanyi*, 1986), ovvero adottando quali valori delle funzioni di utilità le medie aritmetiche dei valori elicitati con le interviste ai singoli *decision maker* coinvolti.

In pratica, il metodo prevede che dato un set di indicatori, con caratteristiche definite nel prosieguo, si associ ad ogni indicatore una funzione di Utilità e quindi, attraverso una funzione matematica definita Utilità Multi-Attributo, si compongano esprimendo così l’utilità $U(\bar{X})$ di ogni possibile *performance* $\bar{X} = (x_1, x_2, x_3, \dots, x_i, \dots, x_n)$ (*Keeney and Raiffa*, 1976; *Farquhar*, 1984).

Figura 5b.2 - Diagramma di flusso metodologia per la definizione di un indice di Performance dei Sistemi Sanitari



Per delineare le dimensioni della *performance* dei sistemi sanitari si è adottato l'approccio delle *Balanced Scorecard* (Kaplan e Norton, 1992), anche in considerazione del loro già ampio utilizzo nella sanità pubblica.

La natura pubblica dei sistemi sanitari ha suggerito l'opportunità di integrare le classiche prospettive proposte da Kaplan e Norton con un'ulteriore prospettiva, rappresentata dall'equità (Fioravanti e Spandonaro, 2004); quindi le dimensioni delle *performance* adottate (fra parentesi la specificazione di alcune classi di variabili afferenti ad ogni dimensione), sono:

- Economico-Finanziaria (efficienza ed economicità, etc.);
- Cliente (qualità percepita, fidelizzazione etc.);
- Innovazione e apprendimento (qualità, appropriatezza clinica, etc.);
- Business interno (appropriatezza organizzativa, integrazione dei servizi, etc.);
- Equità (di accesso, *burden* finanziario sulle famiglie, etc.).

Affinché il modello possa rappresentare una coerente rappresentazione della struttura di preferenza sociale, è necessario che gli indicatori rappresentativi delle diverse dimensioni siano completi (ossia che coprano tutti gli aspetti del processo decisionale), importanti (ossia rappresentativi di un criterio significativo nel processo decisionale), misurabili, familiari, non ridondanti, indipendenti, etc.

Al fine di ottemperare ai criteri sopra descritti è stato individuato un primo set di 11 indicatori per simulare gli effetti del metodo.

Tabella 5b.1 - Indicatori e dimensioni di appartenenza

DIMENSIONE	INDICATORE
Economico-Finanziaria (efficienza, economicità, etc.)	% di disavanzo
	Spesa sanitaria totale (pubblica + 30% privata) pro-capite
Business interno (appropriatezza, qualità organizzativa, integrazione, etc.)	Tasso di ricovero per BPCO, diabete e scompenso cardiaco per residenti 50-74 anni
	Tasso standardizzato di ospedalizzazione per acuti in regime ordinario
	% ricoveri ospedalieri con DRG inappropriati
Cliente (qualità percepita, fidelizzazione etc.)	Tasso di persone in buona salute
	Persone molto soddisfatte dall'assistenza medica ospedaliera
Innovazione	% Procedure chirurgiche in artroscopia e laparoscopia sul totale dei DRG chirurgici
	Quota spesa per la formazione sul totale
Sociale (Equità)	% Famiglie impoverite
	% famiglie con spese catastrofiche

L'indipendenza è condizione assiomatica ai fini dell'implementabilità del metodo; la non ridondanza è altrettanto importante, in quanto il non rispetto comporta la duplicazione delle preferenze, distorcendo i risultati.

In prospettiva ci sarebbe l'intenzione di demandare anche la scelta degli indicatori ad una consultazione fra gli *stakeholder*, chiedendo loro di individuare i più rappresentativi da un set pre-selezionato.

5b.4 Una simulazione della valutazione della performance dei Sistemi Sanitari Regionali

Al fine di testare la sensibilità dell'algoritmo sono state predisposte tre simulazioni, ipotizzando in prima battuta che il valore attribuito agli indicatori decresca linearmente dalla *best performance* alla *worst performance* e anche che il loro contributo sia paritario all'interno della dimensione di afferenza.

Dopo di che si sono formulati tre scenari:

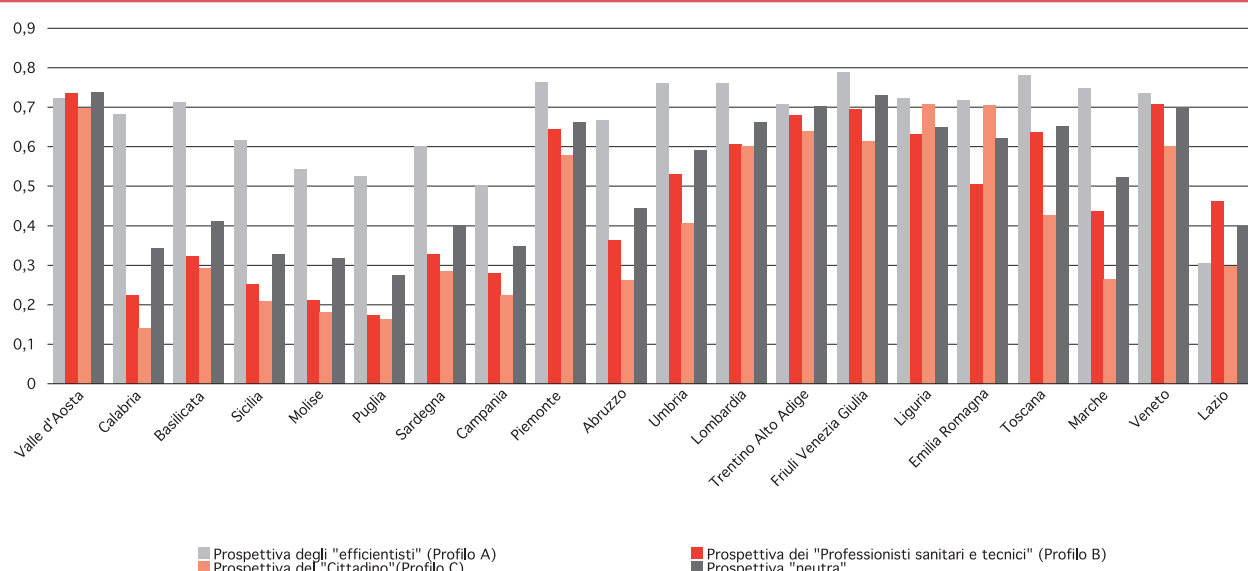
1. nel primo tutte le dimensioni contribuiscono in egual modo alla valutazione della performance complessiva;
2. nel secondo sono stati definiti tre profili di *stakeholder*; il profilo A, che si assume essere efficientista e che dia quindi maggior peso alla dimensione economico-finanziaria (il 60%, lasciando il 40% equidistribuito tra le restanti); il profilo B, rappresentativo dei professionisti sanitari e dei "tecnici", privilegia gli aspetti di organizzazione sanitaria (31%), di 'innovazione' ed 'equità' (circa il 25%), poi la dimensione 'cliente' (17%), lasciando un contributo bassissimo a quella economico-finanziaria (circa il 3%); il profilo C, rappresentativo dei cittadini, invece attribuisce alla dimensione 'cliente' un peso predominante (70%), quasi trascurabile a quella 'economico-finanziaria';
3. nel terzo scenario la determinazione dell'indice unico di *performance* avviene in base alle

preferenze in media espresse dai tre profili di *stakeholder* del Sistema sopra descritti. Evidenziando che si tratta di mere simulazioni, tese a testare la sensibilità del metodo e non “classifiche” regionali, nel grafico che segue è possibile apprezzare come, con gli indicatori proposti, le Regioni del Nord e le Province Autonome presentino mediamente un valore di *performance* significativamente maggiore rispetto a quello delle Regioni del Centro e del Mezzogiorno. Al variare del “peso” attribuito alle diverse dimensioni si modifica però l'indice di *performance* dei SSR e quindi il *ranking*. Attribuendo un maggior peso alla dimensione ‘economico-finanziaria’ (profilo A), rispetto allo scenario base di equidistribuzione dell'importanza delle dimensioni, è possibile notare come la Toscana, ad esempio, passi dal settimo al secondo posto, il Piemonte dal quinto al terzo, l'Umbria dal decimo al quarto, la Puglia dall'ultimo al terz'ultimo, la Campania perde tre posizioni diventando penultima, il Lazio passa dal quindicesimo al ventesimo.

Attribuendo un maggior peso invece alle dimensioni “tecniche” (profilo B) è possibile notare come, ad esempio, il Veneto passi dalla quarta alla seconda posizione, il Lazio recuperi quattro posizioni, passando dalla quindicesima all'undicesima.

Attribuendo un maggior peso invece alla dimensione ‘cliente’ (profilo C) si nota come Regioni quali la Liguria, l'Emilia Romagna e il Lazio modifichino in maniera significativa la loro posizione passando rispettivamente dall'ottava alla prima, dalla nona alla seconda, e dalla quindicesima all'undicesima; la Calabria perde ancora una posizione, diventando ultima.

Figura 5b.3 - Valore di *performance* dei SSR al variare delle prospettive



I *ranking* ottenuti, seppure nell'ambito di un mero esercizio di simulazione, appaiono convincenti, sia in modo aggregato, sia nelle specificazioni che assumono variando l'importanza degli obiettivi da perseguire.

Come era lecito aspettarsi, e a giustificazione dell'importanza di utilizzare un metodo democratico nella elicitazione dei pesi attribuiti ai vari obiettivi/dimensioni, le differenze dei *ranking* possono essere significative, a riprova della complessità del settore sanitario, ma anche a giustificazione della importanza di valutare complessivamente le politiche sanitarie, che possono avere impostazioni legittimamente diverse e quindi risultati non sovrapponibili su singole dimensioni di *performance*.

5b.5 Discussione

Le simulazioni riportate, pur con i limiti preliminarmente espressi, permettono di apprezzare la relativa semplicità di implementazione del metodo, la cui maggiore complessità rimane quella della elicitazione (democratica) delle preferenze, attività peraltro di grande utilità, in quanto fondamentale elemento di conoscenza e guida delle politiche sanitarie, sembrano dimostrare l'importanza di perseguire la massima trasparenza nell'implementazione delle valutazioni.

Le simulazioni hanno consentito di apprezzare la sensibilità dell'algoritmo al diverso "peso" attribuito alle dimensioni della *performance*, senza peraltro produrre stravolgimenti nel *ranking*: i risultati sembrano comunque convincenti e ragionevoli.

L'esercizio "dimostra", in ultima istanza, che l'assunzione di pesi uguali per ogni dimensione sia tutt'altro che scontata: la neutralità assoluta appare poco probabile; le preferenze elicitabili da diverse categorie di *stakeholder* rendono la classifica dei SSR assolutamente non "arbitraria", seppure costruita su preferenze soggettive.

La metodologia proposta risulta relativamente innovativa, quanto meno nella sua applicazione sanitaria; avvalendosi di un ridotto numero di elementi e indicatori (per le motivazioni sopra esposte), il metodo permette di giungere ad una valutazione della *performance* dei Sistemi Sanitari basandosi su un fondamento teorico (la teoria dell'utilità ampiamente utilizzata nelle scienze sociali), rispettando le preferenze soggettive degli *stakeholder*, eliminando i rischi di arbitrarietà insiti nella definizione di pesi prescelti dal "gestore del modello" di valutazione o, peggio, di scarsa trasparenza derivante dalla non esplicitazione dei pesi stessi; è inoltre semplice e flessibile, in quanto consente di implementare diversi criteri per la determinazione dei valori "ottimali" e "peggiori"; lo stesso algoritmo può essere utilizzato per effettuare valutazioni e monitoraggi della *performance* a diversi livelli: aziendale, regionale, nazionale, etc.; infine, la sua applicazione ad una o più dimensioni può consentire di effettuare una valutazione della *performance* parziale relativa alla/e dimensione/i considerata/e.