



UNIVERSITA' POLITECNICA DELLE MARCHE
FACOLTA' DI MEDICINA E CHIRURGIA

Corso di Laurea Magistrale in Scienze Infermieristiche ed
Ostetriche

**“Progettazione e sperimentazione di un sistema di
valutazione delle performance per i valutatori GAAR”**

**“Design and Pilot Testing of a Performance Evaluation
System for GAAR Evaluators”**

Relatore:

**Prof. Domenico
Lombardi Fortino**

Tesi di Laurea:

Matteo Barabani

Correlatore

Dott.ssa Sonia Tonucci

A.A 2025/2026

INDICE

Abstract Italiano.....	1
Abstract in English.....	4
INTRODUZIONE	7
AUTORIZZAZIONE ALL'ESERCIZIO E ACCREDITAMENTO ISTITUZIONALE DELLE STRUTTURE SANITARIE E SOCIO SANITARIE.	7
CONCETTO GENERALE DI VALUTAZIONE.....	10
ACCOUNTABILITY E VALUTAZIONE DEL PERSONALE SANITARIO	11
SISTEMA DI VALUTAZIONE DELLE PERFORMANCE DEI GAAR: STATO ATTUALE	12
OBIETTIVO	18
MATERIALI E METODI	18
1° FASE: REVISIONE DELLA LETTERATURA	19
FASE PRELIMINARE	19
SCOPING REVIEW	20
2° FASE: VALIDAZIONE DELLA SCHEDA DI VALUTAZIONE.....	29
FOCUS GROUP PER LA COSTRUZIONE E VALIDAZIONE PRELIMINARE DELLA SCHEDA DI VALUTAZIONE	29
LA SCHEDA DI VALUTAZIONE "MILLER - AGENAS".....	31
VALIDAZIONE PRELIMINARE DELLA NUOVA SCHEDA DI VALUTAZIONE "MILLER AGENAS".	34
QUESTIONARIO PER LA VALUTAZIONE PERCETTIVA DELLA SCHEDA DI VALUTAZIONE "MILLER AGENAS"	36
METODOLOGIA DI ANALISI DEI DATI QUALITATIVI OTTENUTI	39
METODI STATISTICI.....	40
RISULTATI.....	42

DISTRIBUZIONE DELLE RISPOSTE AI QUESITI DEL QUESTIONARIO DI GRADIMENTO	42
ANALISI STATISTICA	46
DISCUSSIONE	55
SISTEMA DI MISURAZIONE E VALUTAZIONE DELLE PERFORMANCE	
AGENAS – Agenzia nazionale per i servizi sanitari regionali	55
LA PIRAMIDE DI MILLER	57
VALUTAZIONE CRITICA DEL SISTEMA PRECEDENTEMENTE IN USO	59
SIGNIFICATO METODOLOGICO DELL'INTEGRAZIONE TRA PIRAMIDE DI MILLER E SISTEMA AGENAS	61
INTERPRETAZIONE DEI RISULTATI DEL QUESTIONARIO DI GRADIMENTO	61
Grafico 1: Chiarezza del sistema di valutazione	62
Grafico 2: Utilità della valutazione nelle diverse fasi (pre-audit, audit, conclusione).	63
Grafico 3: Chiarezza dei criteri di valutazione	63
Grafico 4: Pertinenza della dimensione partecipativa/collaborativa	63
Grafico 5: Contributo al miglioramento professionale.....	64
Grafico 6: Familiarità con la Piramide di Miller	64
Grafico 7: Familiarità con il sistema AGENAS	64
INTERPRETAZIONE COMPLESSIVA DEI RISULTATI	65
IMPLICAZIONI PER IL SISTEMA DI AUTORIZZAZIONE E ACCREDITAMENTO.....	68
PUNTI DI FORZA DELLO STUDIO	68
LIMITI METODOLOGICI DELLO STUDIO	69
PROSPETTIVE DI SVILUPPO	69
CONCLUSIONI	71

Bibliografia	74
Allegati.....	78

Indice delle figure

Figura 1 Prisma 2020 Diagram	22
Figura 2 Cronoprogramma delle fasi metodologiche	41
Figura 3 Grafico 1 – Frequenza risposta al quesito 1.	42
Figura 4 Grafico 2 - Frequenza risposta al quesito 2.	43
Figura 5 Grafico 3- Frequenza risposta al quesito 3.	43
Figura 6 Grafico 4 - Frequenza risposta al quesito 4.	44
Figura 7 Grafico 5 - Frequenza risposta al quesito 5.	44
Figura 8 Grafico 6 - Frequenza risposta al quesito 6.	45
Figura 9 Grafico 7 - Frequenza risposta al quesito 7.	45
Figura 10 Media degli item Likert con intervallo di confidenza al 95%.....	48
Figura 11 Distribuzione dello score composito di percezione.....	49
Figura 12 Distribuzione della familiarità con Piramide di Miller e sistema AGENAS .	51
Figura 13 La Piramide di Miller	59

Indice tabelle

Tabella 1 Scala di Valutazione dei GAAR	15
Tabella 2 Mappatura degli Item della Scheda di Valutazione GAAR.....	17
Tabella 3 Metodologia PICO.....	19
Tabella 4 Tabella Riassuntiva delle fonti bibliografiche	23
Tabella 5 Summary Of Findings "SOF"	24
Tabella 6 Statistiche descrittive degli item Q1-Q5	47
Tabella 7 Distribuzione percentuale delle risposte agli item Likert	47
Tabella 8 Test di Wilcoxon degli item rispetto al valore neutro della scala (3=abbastanza)	49
Tabella 9 Consistenza interna dello strumento	50
Tabella 10 Analisi fattoriale esplorativa: carichi sul primo fattore	50
Tabella 11 Matrice di correlazione Spearman tra item Q1-Q5.....	51

Tabella 12 Score composito per familiarità con la Piramide di Miller.....	52
Tabella 13 Score composito per familiarità con il sistema AGENAS.....	52
Tabella 14 Associazione tra familiarità con Piramide di Miller e sistema AGENAS	52
Tabella 15 Regressione lineare multipla sullo score composito di percezione	53
Tabella 16 Analisi tematica delle risposte aperte	54

Abstract Italiano

Background

La valutazione, intesa come processo sistematico di misurazione, analisi e interpretazione delle performance, rappresenta uno strumento essenziale nei sistemi organizzativi complessi, in quanto consente di verificare il raggiungimento di standard, obiettivi e livelli di competenza attesi. In ambito sanitario, essa assume una funzione strategica di garanzia della qualità, promozione del miglioramento continuo e supporto alla trasparenza dei processi decisionali. La presenza di criteri strutturati, indicatori osservabili e modelli teorici di riferimento costituisce un elemento fondamentale per assicurare uniformità, oggettività e affidabilità nella valutazione delle performance professionali. Nel contesto dei processi di autorizzazione e accreditamento sanitario, la valutazione delle competenze dei valutatori rappresenta un elemento strategico per garantire qualità, trasparenza e affidabilità del sistema. Tuttavia, dalla letteratura scientifica e dalle fonti istituzionali emerge l'assenza di modelli specificamente sviluppati per la valutazione dei valutatori del Gruppo di Autorizzazione e Accreditamento Regionale (GAAR).

Obiettivo

Individuare modelli teorici e sistemi di valutazione delle competenze applicabili al contesto dei valutatori GAAR e di sviluppare, implementare e valutare un nuovo strumento valutativo strutturato.

Materiali e Metodi

Lo studio è stato sviluppato secondo un approccio multi-metodo, articolato in due fasi integrate e complementari. La prima fase ha previsto la conduzione di una scoping review, finalizzata a esplorare sistematicamente i modelli, gli strumenti e gli indicatori utilizzati per la valutazione delle performance in ambito sanitario. Il processo di identificazione, selezione e inclusione degli studi è stato strutturato secondo il framework PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), al fine di garantire trasparenza metodologica, tracciabilità del processo decisionale e rigore scientifico nella mappatura delle evidenze disponibili. L'analisi critica degli articoli inclusi ha consentito di individuare dimensioni valutative ricorrenti e di criteri di

misurazione delle performance ritenuti maggiormente significativi in letteratura, costituendo la base teorico-metodologica per la seconda fase dello studio. In particolare, le evidenze emerse hanno supportato la progettazione e l'avvio del processo di validazione di una scheda di valutazione delle performance, costruita a partire dall'integrazione di due modelli concettuali di riferimento ampiamente riconosciuti in letteratura scientifica. La fase di sperimentazione dello strumento è stata accompagnata dalla somministrazione di un questionario di gradimento rivolto ai valutatori coinvolti nel processo applicativo.

Risultati

Il processo di valutazione delle performance dei GAAR iscritto nell'elenco regionale della Regione Marche relativo all'anno 2025 è stato condotto mediante l'impiego della scheda sperimentale su una popolazione complessiva di 115 soggetti appartenenti al gruppo dei valutatori GAAR. La sperimentazione dello strumento è stata valutata mediante questionario di gradimento somministrato ai valutatori GAAR. Il dataset analizzato comprendeva 73 rispondenti, pari al 63.5% della popolazione eleggibile ($n=115$). I primi cinque item, misurati su scala Likert 1–5, hanno mostrato punteggi medi elevati (range: 4,32–4,44), con una quota di risposte positive (Molto/Completamente) compresa tra l'82,2% e il 90,4%. Lo score composito di percezione è risultato pari a 4.36 ± 0.68 (IC95% 4.20–4.52) e significativamente superiore al valore neutro della scala (Wilcoxon $p < 0,001$). La consistenza interna dello strumento è risultata eccellente (α di Cronbach=0.900) e l'analisi fattoriale esplorativa ha evidenziato una struttura monodimensionale, con primo fattore responsabile del 71.9% della varianza. Le analisi bivariate e multivariate non hanno evidenziato associazioni statisticamente significative tra familiarità pregressa con Piramide di Miller/sistema AGENAS e score complessivo. I risultati supportano l'accettabilità e la validità applicativa preliminare del nuovo sistema valutativo.

Conclusioni

Il modello “Miller–AGENAS” sviluppato si configura come uno strumento metodologicamente solido, caratterizzato da elevata accettabilità, eccellente coerenza interna e preliminare validità applicativa nel contesto della valutazione dei valutatori GAAR. L’integrazione tra framework competence-based e sistema istituzionale di misurazione della performance ha consentito la costruzione di un sistema valutativo strutturato, potenzialmente trasferibile ad altri contesti organizzativi complessi. Ulteriori studi su campioni più ampi e in differenti setting operativi risultano necessari per confermarne stabilità psicometrica, validità esterna e sostenibilità nell’implementazione istituzionale su larga scala.

Parole chiave: GAAR, OTA, Valutazione delle performance, Piramide di Miller, AGENAS, Validazione di strumenti, Competenze professionali.

Abstract in English

Background

Evaluation, understood as a systematic process of measuring, analyzing, and interpreting performance, represents an essential tool within complex organizational systems, as it enables the verification of standards achievement, objectives attainment, and expected competency levels. In healthcare settings, it assumes a strategic role in ensuring quality, promoting continuous improvement, and supporting transparency in decision-making processes. The presence of structured criteria, observable indicators, and theoretical reference models therefore constitutes a fundamental element in ensuring uniformity, objectivity, and reliability in the assessment of professional performance. Within healthcare authorization and accreditation processes, the assessment of assessors' competencies represents a strategic element to ensure system quality, transparency, and reliability. However, both the scientific literature and institutional sources highlight the absence of models specifically developed for evaluating members of the Regional Authorization and Accreditation Group (GAAR).

Objective

To identify theoretical models and competency assessment systems applicable to the GAAR context, and to develop, implement, and evaluate a new structured assessment tool.

Material and methods

The study was developed according to a multi-method approach, structured into two integrated and complementary phases. The first phase involved the conduction of a scoping review aimed at systematically exploring models, tools, and indicators used for performance evaluation in healthcare settings. The process of study identification, selection, and inclusion was structured according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework, in order to ensure methodological transparency, decision-making traceability, and scientific rigor in mapping the available evidence. The critical analysis of the included studies allowed for the identification of recurring evaluative dimensions and performance measurement criteria considered most relevant in the literature, thereby providing the theoretical and methodological foundation for the second phase of the study. Specifically, the evidence

gathered supported the design and initiation of the validation process of a performance evaluation tool, developed through the integration of two conceptual reference models widely recognized in the scientific literature. The experimental phase of the instrument was accompanied by the administration of a satisfaction questionnaire to the assessors involved in its practical application.

Results

The 2025 performance evaluation process of GAAR members registered in the Regional Register of the Marche Region was conducted using the experimental assessment tool on a total population of 115 subjects belonging to the GAAR evaluators group. The experimental implementation of the instrument was assessed through a satisfaction questionnaire administered to GAAR evaluators. The analyzed dataset included 73 respondents, corresponding to 63.5% of the eligible population ($n=115$). The first five items, measured on a 1–5 Likert scale, showed high mean scores (range: 4.32–4.44), with the proportion of positive responses (Very/Completely) ranging from 82.2% to 90.4%. The composite perception score was 4.36 ± 0.68 (95% CI 4.20–4.52) and was significantly higher than the neutral midpoint of the scale (Wilcoxon $p < 0.001$). The internal consistency of the instrument was excellent (Cronbach's $\alpha = 0.900$), and exploratory factor analysis revealed a unidimensional structure, with the first factor accounting for 71.9% of the variance. Bivariate and multivariate analyses did not identify statistically significant associations between prior familiarity with Miller's Pyramid and/or the AGENAS performance system and the overall score. These findings support the acceptability and preliminary applied validity of the new assessment system.

Conclusions

The developed “Miller–AGENAS” model appears to be a methodologically robust tool, characterized by high acceptability, excellent internal consistency, and preliminary applied validity within the context of GAAR evaluators' performance assessment. The integration of a competence-based framework with an institutional performance measurement system enabled the development of a structured assessment model potentially transferable to other complex organizational settings. Further studies involving larger samples and different operational contexts are needed to confirm its

psychometric stability, external validity, and sustainability for large-scale institutional implementation.

Keywords: GAAR, OTA, AGENAS, Performance evaluation, Competency assessment, Miller's Pyramid, Healthcare accreditation, Validation study.

INTRODUZIONE

AUTORIZZAZIONE ALL'ESERCIZIO E ACCREDITAMENTO ISTITUZIONALE DELLE STRUTTURE SANITARIE E SOCIO SANITARIE.

La Regione Marche, in attuazione delle finalità previste dall'art. 1 della Legge Regionale n. 21/2016, disciplina le procedure per il rilascio dell'autorizzazione alla realizzazione, dell'autorizzazione all'esercizio e dell'accreditamento istituzionale delle strutture sanitarie e socio-sanitarie. Per **autorizzazione** si intendono i provvedimenti amministrativi che consentono la realizzazione di strutture e l'esercizio di attività sanitarie e socio-sanitarie da parte di soggetti pubblici e privati, nel rispetto dei requisiti minimi strutturali, tecnologici, impiantistici e organizzativi stabiliti dalla normativa vigente. Per **accreditamento istituzionale** si intende il provvedimento con cui viene riconosciuto alle strutture pubbliche e private già autorizzate lo status di potenziali erogatori di prestazioni per conto del Servizio Sanitario Nazionale (SSN), subordinatamente al rispetto di livelli di qualità ulteriori e più stringenti rispetto a quelli previsti per l'autorizzazione. Autorizzazione e accreditamento sono entrambi processi di valutazione sistematica e periodica finalizzati a verificare il possesso e il mantenimento, da parte delle strutture richiedenti, di requisiti strutturali, tecnologici, impiantistici, organizzativi e funzionali che concorrono a determinare la qualità dell'assistenza erogata. Il sistema regionale di autorizzazione e di accreditamento è disciplinato dalla Legge Regionale n. 21 del 30 settembre 2016 recante "Autorizzazioni e accreditamento istituzionale delle strutture e dei servizi sanitari, socio-sanitari e sociali pubblici e privati, e disciplina degli accordi contrattuali delle strutture e dei servizi sanitari, socio-sanitari e sociali pubblici e privati", successivamente modificata dalla Legge Regionale n. 7 del 14 marzo 2017. A completamento della disciplina il Regolamento Regionale 1° febbraio 2018 n. 1 ha definito e codificato le varie tipologie di strutture e servizi sanitari, socio-sanitari e sociali, in attuazione dell'art. 7, comma 3, della L.R. 21/2016. Il rilascio delle autorizzazioni e degli accreditamenti delle strutture sanitarie e socio-sanitarie extraospedaliere (residenziali e semiresidenziali), nonché delle strutture ospedaliere, ambulatoriali, dei laboratori, degli studi professionali e delle terme, è effettuata dal Settore Autorizzazioni e Accreditamenti, afferente all'Agenzia Regionale Sanitaria della Regione Marche.

Nel contesto del Servizio Sanitario Nazionale, il sistema di autorizzazione e accreditamento istituzionale delle strutture sanitarie rappresenta uno strumento essenziale per garantire qualità, sicurezza e appropriatezza nell'erogazione delle prestazioni. La normativa nazionale, unitamente ai recepimenti e declinazioni regionali ha progressivamente rafforzato l'importanza di un modello di governance fondato sulla responsabilità, la trasparenza e il miglioramento continuo.

La costituzione e il funzionamento di tale organismo sono stati definiti con DGR 1982 del 18 dicembre 2023 recante "Modifica DGR n. 522 del 23/4/2018 L.R. 30 settembre 2016, n. 21 e s.m.i. - Costituzione e funzionamento dell'Organismo Tecnicamente Accreditante (OTA)". Ai sensi dell'intesa Stato Regioni n.32 del 19/02/2015, le Regioni sono tenute ad adeguare il funzionamento dell'OTA alle disposizioni ivi contenute, al fine di uniformare a livello nazionale i processi di autorizzazione e accreditamento delle strutture sanitarie. La costituzione dell'OTA si fonda sui principi di autonomia per l'espletamento delle proprie funzioni, senza conflitto di interessi ed in totale imparzialità e trasparenza nei confronti, sia dell'Organo Amministrativo Accreditante Regionale che delle strutture da accreditare/autorizzare [1]A tal proposito, si richiama quanto previsto dall'articolo 5, comma 1 e 2, della Legge Regionale n. 21/2016, che dispone quanto segue: Comma 1: La Giunta regionale definisce le procedure per l'autorizzazione e l'accREDITamento regionale delle strutture sanitarie e socio-sanitarie, prevedendo, nell'ambito di tali processi, il supporto tecnico-consulativo di un apposito organismo denominato Organismo Tecnicamente Accreditante (OTA). I Comuni possono avvalersi di tale organismo per l'esercizio delle funzioni di propria competenza.

Comma 2: La Giunta regionale stabilisce, in particolare, l'organizzazione e le modalità operative dell'OTA, articolato in un Gruppo centrale e nel Gruppo di Autorizzazione e AccredITamento Regionale (GAAR). Inoltre, determina l'ammontare del contributo istruttorio che i soggetti richiedenti l'autorizzazione o l'accREDITamento devono versare all'Agenzia Regionale Sanitaria (ARS), presso la quale l'OTA è istituito.

Il GAAR, costituito presso l'OTA, è composto da personale degli Enti del Servizio Sanitario Regionale (SSR) che aderisce volontariamente alla selezione. Solo a seguito del superamento della stessa, la Direzione dell'Ente di appartenenza rilascia il nulla osta per la partecipazione alle attività [2]L'inserimento nella lista ufficiale dei GAAR è subordinato al superamento della selezione pubblica per titoli ed esami, finalizzata

all'accesso alla formazione specialistica e, alla valutazione positiva al termine del corso di formazione. I componenti del GAAR sono definiti in uno specifico atto dell'ARS.

Nello specifico il valutatore GAAR è responsabile delle seguenti attività: [3]

- partecipa alla stesura/revisione dei Manuali di Autorizzazione e Accreditamento della Regione Marche;
- valuta il materiale documentale allegato alla domanda preliminarmente alla effettuazione dei sopralluoghi (pre — audit);
- redige il rapporto finale di verifica (verbale) con proposta di rilascio di autorizzazione all'esercizio e accreditamento (fase dell'audit);
- collabora con il Comitato Tecnico Operativo (CTO) per eventuali ulteriori approfondimenti resisi necessari in seguito alla visita ed al rapporto finale di verifica;
- valuta la documentale delle evidenze di tipo III e IV, di cui ai cronoprogrammi forniti dalle Strutture, per i processi di Accreditamento di livello avanzato e di eccellenza;
- valuta la documentazione in materia di accreditamento delle strutture sanitarie e socio sanitarie - Integrazione della DGR 1572/2019;
- valuta l'assolvimento di eventuali prescrizioni rilasciate alla Struttura in sede di verifica;
- effettua visite ispettive, anche in orario di servizio, a seguito di convocazione del Coordinatore dell'Organismo Tecnicamente Accreditante (OTA).

L'attività del GAAR consiste nel verificare, in maniera terza e imparziale, il rispetto dei requisiti di autorizzazione e accreditamento da parte delle strutture sanitarie. Per garantire indipendenza di giudizio ed evitare potenziali conflitti di interesse, i valutatori GAAR della Regione Marche effettuano i sopralluoghi esclusivamente al di fuori della provincia di abituale servizio. Al fine di mantenere elevati standard di trasparenza e affidabilità nel processo, è attivo un sistema di controllo e valutazione periodica delle prestazioni dei valutatori GAAR. In particolare, il Decreto del Dirigente del Settore Territorio ed Integrazione Socio Sanitaria – ARS n. 31 del 17 luglio 2023, recante "Valutazione individuale del gruppo GAAR", definisce in dettaglio le modalità operative per la conduzione di tale valutazione.

CONCETTO GENERALE DI VALUTAZIONE

La valutazione è un processo sistematico, strutturato e documentato finalizzato a misurare, analizzare e interpretare il grado di raggiungimento di obiettivi prestabiliti, la qualità delle prestazioni e l'efficacia dei comportamenti professionali. In ambito organizzativo e, nello specifico, nel contesto sanitario e sociosanitario, essa rappresenta uno strumento strategico di governance e miglioramento continuo, consentendo di monitorare sia la performance individuale sia quella collettiva, in un'ottica di trasparenza e responsabilità (accountability).

Il processo valutativo si fonda su tre elementi cardine:

- Definizione degli obiettivi e degli standard di riferimento: gli obiettivi devono essere chiari, misurabili, temporalmente definiti e coerenti con la mission e la vision dell'organizzazione;
- Raccolta e analisi dei dati: attraverso indicatori quantitativi e qualitativi, strumenti di rilevazione strutturati (schede, check-list, griglie di osservazione) e, ove possibile, il confronto con benchmark interni o esterni;
- Attribuzione del giudizio: basata su criteri oggettivi, condivisi e verificabili, spesso supportata da scale di misurazione (es. Likert) che consentono di quantificare il livello di performance.

La valutazione ha una duplice funzione:

- Funzione di controllo: verifica la conformità ai requisiti previsti da norme, regolamenti e standard organizzativi;
- Funzione di sviluppo: individua punti di forza e aree di miglioramento, favorendo interventi formativi, correttivi e innovativi.

Un sistema di valutazione efficace deve rispettare alcuni principi fondamentali:

- Oggettività: criteri e metodi chiari, misurabili e replicabili;
- Trasparenza: conoscenza preventiva da parte dei valutati delle modalità e degli strumenti utilizzati;
- Imparzialità: assenza di conflitti di interesse e garanzia di equità di giudizio;
- Continuità: svolgimento periodico e regolare, con monitoraggio nel tempo;

- Orientamento al miglioramento: utilizzo dei risultati per pianificare azioni correttive e potenziare le competenze.

Nel contesto della Pubblica Amministrazione italiana, la valutazione è inquadrata dal **D.Lgs. n. 150/2009**, che ha introdotto il “*ciclo di gestione della performance*” e stabilito che la misurazione e la valutazione devono riguardare:

- Il raggiungimento degli obiettivi organizzativi;
- La qualità e la tempestività delle prestazioni;
- I comportamenti professionali e organizzativi;

Il contributo individuale e di gruppo al miglioramento complessivo del servizio.

ACCOUNTABILITY E VALUTAZIONE DEL PERSONALE SANITARIO

Nel settore pubblico, il concetto di accountability rappresenta un principio cardine della legittimazione dell'azione amministrativa, configurandosi come l'obbligo, in capo ai soggetti titolari di funzioni pubbliche, di rendere conto delle decisioni assunte, delle risorse impiegate e dei risultati conseguiti nei confronti dei cittadini e degli altri stakeholders [4]. In ambito sanitario, tale principio assume una rilevanza ancora maggiore in considerazione della natura costituzionalmente tutelata del diritto alla salute e dell'ingente quantità di risorse pubbliche destinate al Servizio Sanitario Nazionale (SSN). Come evidenziato dalla letteratura, l'accountability non si esaurisce in una dimensione meramente contabile o finanziaria, ma si articola in un processo sistematico di misurazione, valutazione e rendicontazione delle performance [4]. In questo senso, la valutazione costituisce lo strumento attraverso cui l'accountability prende forma operativa, consentendo di verificare la coerenza tra obiettivi programmati e risultati effettivamente raggiunti, nonché di rafforzare la trasparenza e la credibilità dell'azione pubblica [4]. Nel contesto del SSN italiano, l'evoluzione normativa e organizzativa ha progressivamente spostato l'attenzione da un'accountability di tipo formale e burocratico ad una accountability di tipo manageriale e professionale, centrata sulla qualità delle prestazioni, sull'efficacia clinica, sull'appropriatezza e sulla sicurezza delle cure [4]. Tale transizione è coerente con il passaggio da un modello di controllo prevalentemente finanziario ad un approccio multidimensionale alla performance sanitaria, che integra indicatori economici con indicatori di esito, processo ed equità [5]. In questo quadro, la

valutazione del personale sanitario e, più in generale, dei soggetti che esercitano funzioni di indirizzo, controllo e verifica, si configura come un presupposto essenziale per garantire l'effettività del principio di accountability. Se infatti l'accountability implica la responsabilità di rendere conto dell'operato e dei risultati conseguiti, essa non può prescindere dall'esistenza di strumenti strutturati e oggettivabili di valutazione delle competenze e delle performance professionali [6]. Particolare rilievo assume, in tal senso, la dimensione multilivello dell'accountability nel SSN, che si articola in forme verticali (nei confronti dei cittadini) e orizzontali o interistituzionali (tra diversi livelli di governo e tra enti controllori e controllati) [7]. All'interno di tale architettura, i soggetti incaricati di verificare la qualità e la conformità delle strutture sanitarie, **quali i gruppi di valutazione per l'autorizzazione e l'accreditamento**, non solo esercitano una funzione di controllo, ma sono a loro volta inseriti in un circuito di responsabilità che richiede coerenza, trasparenza e tracciabilità delle proprie decisioni. Ne deriva che la costruzione di sistemi di valutazione strutturati, fondati su criteri espliciti e condivisi, rappresenta non soltanto uno strumento di miglioramento organizzativo, ma un elemento imprescindibile per il rafforzamento della legittimazione istituzionale e della fiducia nei confronti dell'intero sistema sanitario. In questa prospettiva, la valutazione del personale sanitario e dei valutatori stessi si colloca all'interno di una più ampia logica di accountability pubblica, orientata alla qualità, alla trasparenza e alla sostenibilità del Servizio Sanitario Nazionale.

SISTEMA DI VALUTAZIONE DELLE PERFORMANCE DEI GAAR: STATO ATTUALE

L'introduzione di un sistema di valutazione delle performance rivolto ai GAAR risponde alla necessità di garantire uniformità, trasparenza e tracciabilità nei processi di autorizzazione e accreditamento delle strutture sanitarie e socio-sanitarie. La crescente complessità del sistema regionale di accreditamento, caratterizzato da elevati livelli di responsabilità tecnica, multidisciplinarietà e rilevanza istituzionale, richiede infatti strumenti strutturati in grado di monitorare in maniera sistematica le competenze professionali, la qualità delle attività svolte e la coerenza operativa rispetto agli standard regionali e nazionali di riferimento. L'adozione di un sistema di valutazione consente inoltre di rafforzare i principi di accountability professionale e organizzativa,

promuovendo processi decisionali basati su criteri oggettivi, osservabili e misurabili. In tale prospettiva, la valutazione rappresenta non solo uno strumento di verifica delle performance individuali, ma anche un dispositivo di governance orientato al miglioramento continuo della qualità dei processi valutativi e dell'affidabilità complessiva del sistema di accreditamento istituzionale. Ulteriore elemento di rilevanza riguarda la necessità di standardizzare i comportamenti valutativi dei professionisti coinvolti nelle attività di audit e verifica, riducendo la variabilità interpretativa e favorendo una maggiore omogeneità nell'applicazione dei requisiti di autorizzazione e accreditamento. La presenza di criteri valutativi condivisi contribuisce infatti a incrementare la riproducibilità delle valutazioni, la comparabilità degli esiti e la credibilità istituzionale del sistema regionale. L'implementazione di un sistema strutturato di valutazione delle performance dei GAAR assume inoltre una funzione strategica nell'identificazione dei bisogni formativi e nello sviluppo delle competenze professionali avanzate. Attraverso l'analisi sistematica delle aree di forza e delle criticità emergenti, il sistema può supportare la programmazione di interventi formativi mirati, coerenti con l'evoluzione normativa, organizzativa e metodologica del settore. Infine, l'applicazione di un modello valutativo formalizzato risulta coerente con gli orientamenti nazionali in materia di misurazione e valutazione della performance nella pubblica amministrazione e nei sistemi sanitari, favorendo l'integrazione tra qualità, sicurezza, efficacia organizzativa e trasparenza amministrativa. La necessità di sottoporre i valutatori GAAR a un processo di valutazione strutturato deriva anche dalla peculiare posizione che essi occupano all'interno dell'architettura del sistema di autorizzazione e accreditamento: i GAAR, infatti, esercitano una funzione tecnica di verifica esterna che incide direttamente sulla legittimazione, sul mantenimento e sul miglioramento degli standard qualitativi delle strutture sanitarie e socio-sanitarie regionali. In tale prospettiva, la qualità del sistema valutativo rivolto alle strutture non può prescindere dalla qualità, dall'affidabilità e dalla coerenza metodologica dei professionisti chiamati a esercitare tale funzione. La valutazione dei GAAR risponde pertanto all'esigenza di garantire che le competenze richieste non siano considerate acquisite in modo statico, ma siano oggetto di monitoraggio periodico, aggiornamento e verifica continua, in relazione all'evoluzione normativa, organizzativa e tecnico-professionale del settore. Inoltre, la natura stessa dell'attività di audit, fondata su osservazione, interpretazione applicativa dei requisiti e

formulazione di giudizi tecnici, rende essenziale disporre di strumenti in grado di presidiare il rischio di disomogeneità valutativa, discrezionalità eccessiva o progressivo scostamento dagli standard metodologici condivisi. In questo senso, la valutazione periodica dei GAAR assume una funzione di garanzia sistemica, in quanto contribuisce a preservare la coerenza interna dell'intero processo di autorizzazione e accreditamento, rafforzando la qualità tecnica delle verifiche e la solidità del sistema regionale nel suo complesso. Essa si configura, dunque, non soltanto come verifica delle performance individuali, ma come elemento essenziale di manutenzione metodologica e istituzionale di un sistema ad alta complessità organizzativa [8]. La valutazione è condotta con cadenza annuale nel mese di dicembre e si riferisce all'anno solare appena concluso, sotto la responsabilità del Coordinatore OTA. L'attuale scheda di valutazione GAAR impiega un sistema strutturato, articolato in specifici items di performance, ciascuno associato a un punteggio numerico e a una corrispondente classificazione qualitativa.

Struttura della scheda e items valutati (allegato 1)

Gli items includono diverse aree di attività del valutatore GAAR, in particolare riguardano:

- a) Effettuazione delle visite ispettive coordinate dall'OTA – disponibilità e partecipazione alle visite nelle strutture sanitarie e socio-sanitarie.
- b) Accuratezza nella redazione dei verbali di visita ispettiva – precisione e rispetto della tempistica (entro 5 giorni).
- c) Esecuzione del pre-audit e accuratezza nella compilazione del “Piano di Audit” – qualità nella preparazione e documentazione pre-visita.
- d) Accuratezza nella redazione dei verbali per autorizzazione e accreditamento – rispetto della tempistica (entro 10 giorni).
- e) Accuratezza nella valutazione di assolvimento delle prescrizioni – verifica puntuale delle evidenze documentali e tempi di trasmissione (entro 10 giorni).
- f) Accuratezza nella valutazione delle evidenze documentali di accreditamento – anche per accreditamento avanzato/eccellenza (entro 15 giorni).
- g) Partecipazione a eventi formativi specifici GAAR – impegno nell'aggiornamento professionale.

- h) Partecipazione alle riunioni indette dall'OTA – presenza e coinvolgimento in riunioni operative.
- i) Disponibilità a collaborare nell'individuazione di soluzioni organizzative innovative – capacità di problem solving e proattività.

Il sistema di valutazione adottato prevede l'assegnazione di punteggi secondo una scala Likert (1-10), con la seguente graduazione per ciascun item:

Tabella 1 Scala di Valutazione dei GAAR

Giudizio qualitativo	Valore Likert (singolo item)	Punteggio complessivo (giudizio finale)
Eccellente	10	80 – 90
Buono	8 – 9	66 – 79
Sufficiente	6 – 7	54 – 65

Qualora, per uno specifico item, non siano state assegnate dall'OTA attività corrispondenti, viene attribuito il punteggio massimo previsto.

Il sistema di valutazione prevede, nella fase conclusiva, un momento di confronto strutturato tra valutatore e valutato, finalizzato a definire strategie di sviluppo e di miglioramento delle performance, con particolare attenzione ai casi in cui siano stati attribuiti punteggi inferiori agli standard attesi. In un'ottica di trasparenza e tracciabilità, la scheda di valutazione include sezioni dedicate alla rilevazione delle criticità emerse, all'individuazione delle azioni di miglioramento concordate, alla condivisione del punteggio assegnato e alla possibilità per il valutato di accettare o meno il giudizio espresso. In caso di non accettazione, il valutato può inserire annotazioni a supporto della propria posizione. Il sistema integra, inoltre, ulteriori indicatori che consentono di completare e affinare il giudizio finale, garantendo una valutazione più oggettiva e multidimensionale.

Scheda di valutazione del gruppo di audit GAAR da parte delle strutture sanitarie visitate (allegato 2)

Questa Scheda è uno strumento di rilevazione strutturato predisposto dall'OTA, volto a raccogliere feedback qualitativi e quantitativi in merito alla performance del team di

valutatori durante le visite di autorizzazione e accreditamento istituzionale. Dal punto di vista metodologico, la scheda è articolata in tre sezioni principali, ognuna finalizzata a osservare specifiche dimensioni della performance valutativa. La prima sezione riguarda la percezione complessiva del team di valutazione, analizzando indicatori trasversali quali la puntualità, il rispetto del programma operativo, la capacità di affrontare e risolvere situazioni critiche, il livello di competenza tecnica dimostrata, la coesione interna del gruppo, la riservatezza nella gestione delle informazioni e la chiarezza nella condivisione delle eventuali prescrizioni con la struttura valutata. La seconda sezione si concentra sul Team Leader, valutandone le competenze di leadership e gestione: capacità di instaurare un clima collaborativo e costruttivo, puntualità, abilità di coordinamento e autorevolezza, rispetto della pianificazione concordata e garanzia di riservatezza. La terza sezione prende in esame i singoli componenti del team di valutazione, verificando aspetti quali la partecipazione attiva e propositiva, la puntualità, il rispetto del programma nell'ambito di competenza assegnato, la capacità di mantenere un atteggiamento collaborativo e il rispetto della riservatezza.

Il sistema di misurazione adottato si basa su una scala di valutazione Likert a 10 punti, in cui il punteggio 1 corrisponde al livello minimo di soddisfazione e il punteggio 10 al livello massimo. È inoltre prevista una sezione dedicata a commenti e osservazioni qualitative, utile a contestualizzare i punteggi attribuiti e a fornire indicazioni operative per il miglioramento continuo. Dal punto di vista funzionale, la scheda non si limita a rappresentare un mero strumento di raccolta dati, ma si configura come un elemento integrante del ciclo di qualità del processo di autorizzazione e accreditamento. Essa permette di monitorare la performance del GAAR, rilevare punti di forza e aree di miglioramento, e supportare lo sviluppo di azioni correttive o formative mirate. In questo senso, il feedback sistematico fornito dalle strutture valutate costituisce un indicatore indiretto ma strategico dell'efficacia complessiva del sistema di accreditamento regionale.

Tabella 2 Mappatura degli Item della Scheda di Valutazione GAAR

ITEM	AREA DI INDAGINE	OBIETTIVO VALUTATIVO
Puntualità del team	Organizzazione e rispetto tempi	Verificare l'affidabilità e la capacità di gestione del tempo del gruppo di valutazione
Rispetto del programma	Gestione operativa	Assicurare l'aderenza al piano di audit e agli obiettivi della visita
Gestione di situazioni critiche	Problem solving	Valutare la prontezza e l'efficacia nella risoluzione di problemi imprevisti
Competenza specifica	Competenze tecniche	Misurare il livello di preparazione e padronanza della materia
Coesione operativa del gruppo	Team working	Verificare la capacità di lavorare in modo coordinato e sinergico
Riservatezza	Etica e professionalità	Garantire il rispetto della privacy e la tutela delle informazioni
Condivisione delle prescrizioni	Comunicazione e trasparenza	Assicurare chiarezza e completezza nelle informazioni trasmesse alla struttura
Capacità di instaurare un clima collaborativo (Team Leader)	Leadership e gestione relazioni	Favorire un ambiente di lavoro costruttivo e aperto al dialogo
Coordinamento e autorevolezza (Team Leader)	Leadership	Verificare la capacità di guidare il gruppo e mantenere l'orientamento agli obiettivi
Partecipazione attiva e propositiva (Valutatori)	Coinvolgimento operativo	Misurare il grado di contributo individuale durante l'audit

Il valutatore GAAR è sottoposto a una valutazione specifica anche in relazione ad alcuni elementi interni, con particolare attenzione alla verifica dell'effettivo svolgimento delle seguenti attività: Contributo alla stesura, revisione e aggiornamento dei Manuali di Autorizzazione e Accredimento della Regione Marche;

Il sistema di valutazione assicura la qualità e la coerenza dell'operato dei valutatori GAAR, garantendo al contempo il rispetto dei principi di imparzialità, competenza e trasparenza nel processo di accreditamento istituzionale. In questa prospettiva, la valutazione dei valutatori GAAR si configura come un nodo strategico all'interno del più ampio sistema di autorizzazione e accreditamento, rappresentandone un elemento chiave di affidabilità e credibilità. L'adozione di criteri valutativi aggiornati, strutturati e allineati agli standard di qualità definiti dall'OTA, supportati da un solido framework basato su evidenze scientifiche consolidate, rappresenta non solo una necessità tecnico-operativa, ma anche una garanzia di efficacia e trasparenza per l'intero processo. Tale consapevolezza apre la strada a un'analisi critica delle modalità attualmente adottate, volta a individuare eventuali criticità, margini di miglioramento e soluzioni orientate all'introduzione di strumenti valutativi più oggettivi, standardizzati e replicabili.

OBIETTIVO

Sviluppare un sistema strutturato di valutazione delle competenze dei valutatori GARR e valutarne l'applicabilità, la fattibilità e l'efficacia.

MATERIALI E METODI

Lo studio è stato sviluppato secondo un approccio multi-metodo, articolato in due fasi tra loro strettamente integrate e concepite come parti complementari di un unico percorso metodologico. La prima fase ha previsto la conduzione di una scoping review, finalizzata a esplorare in modo sistematico i modelli teorici, gli strumenti operativi e gli indicatori maggiormente utilizzati nella letteratura scientifica per la valutazione delle performance in ambito sanitario e organizzativo. Il processo di identificazione, selezione e inclusione degli studi è stato strutturato secondo il framework PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [9], al fine di garantire trasparenza metodologica, tracciabilità del processo decisionale e rigore scientifico nella mappatura e nell'analisi delle evidenze disponibili. L'esame critico della letteratura selezionata ha consentito di individuare dimensioni valutative ricorrenti, criteri di misurazione delle performance e modelli concettuali ritenuti particolarmente rilevanti, fornendo così il fondamento teorico-metodologico per la successiva fase progettuale. Sulla base delle evidenze emerse, la seconda fase dello studio è stata orientata allo sviluppo, alla strutturazione e all'avvio del processo di validazione di una nuova scheda di valutazione delle performance, costruita attraverso l'integrazione di due modelli concettuali di riferimento ampiamente riconosciuti in letteratura scientifica e adattati al contesto operativo dei valutatori GAAR. La fase sperimentale dello strumento è stata successivamente accompagnata dalla somministrazione di un questionario di gradimento rivolto ai valutatori coinvolti nel processo applicativo, con l'obiettivo di raccogliere elementi preliminari relativi alla percezione di chiarezza, utilità e applicabilità del nuovo sistema valutativo.

1° FASE: REVISIONE DELLA LETTERATURA

FASE PRELIMINARE

Questa fase ha previsto una revisione narrativa della letteratura scientifica e della normativa vigente, condotta con l'obiettivo di identificare modelli teorici e strumenti valutativi validati applicabili al contesto sanitario e, in particolare, alla valutazione delle competenze professionali. A supporto del processo di ricerca bibliografica, i termini MeSH sono stati identificati mediante l'applicazione della metodologia PICO [10] (tabella 3), che ha consentito di strutturare in modo sistematico i principali elementi concettuali dell'indagine e di derivare le parole chiave (keywords) utilizzate per la costruzione della prima stringa di ricerca. Tale stringa è stata successivamente impiegata per l'interrogazione dei principali database scientifici indicizzati, al fine di garantire coerenza metodologica, pertinenza delle fonti selezionate e maggiore accuratezza nell'individuazione delle evidenze disponibili.

Tabella 3 Metodologia PICO

P	C	C
Operatore Sanitario Operatore Professionale	Strumenti di valutazione Schede di valutazione	Valutazione del valutatore in ambito di accreditamento
Healthcare Worker Professional Worker	Assessment Tools Evaluation Sheets	Evaluator evaluation in the context of accreditation

Tuttavia la ricerca bibliografica non ha consentito di identificare, in maniera immediata, contributi scientifici pienamente coerenti con lo specifico obiettivo dello studio. Questa ricerca si è rivelata però strategica per la definizione di un quadro concettuale più ampio, portando all'individuazione di studi realizzati in contesti professionali eterogenei rispetto a quello sanitario, ma accomunati da un interesse primario per la qualità dell'apprendimento e per le metodologie di valutazione delle competenze. Sebbene tali

contributi non facciano esplicito riferimento al contesto sanitario o ai GAAR, essi si caratterizzano per l'impiego di modelli valutativi strutturati e metodologicamente robusti. Questi modelli hanno immediatamente suscitato interesse, in quanto potenzialmente trasferibili, con gli opportuni adattamenti, all'ambito della valutazione delle competenze cliniche e organizzative oggetto del presente studio, rivolto a un gruppo di professionisti con un livello di expertise avanzato. Attraverso una successiva attività di ricerca delle evidenze secondo metodologia snowball search [11], condotta mediante l'analisi sistematica delle referenze bibliografiche degli studi inizialmente selezionati, è stato individuato un modello teorico di riferimento di elevata rilevanza scientifica: la **“Piramide di Miller”**. Tale modello, ampiamente riconosciuto e validato nell'ambito della formazione sanitaria e della valutazione delle competenze professionali e cliniche, è stato assunto quale principale framework concettuale per la strutturazione metodologica dello studio e per lo sviluppo del sistema di valutazione delle performance.

La Piramide di Miller è stata assunta quale struttura teorica portante del progetto, in quanto consente di articolare il processo valutativo secondo livelli progressivi di acquisizione, integrazione e dimostrazione delle competenze, favorendo un approccio multidimensionale orientato alla misurazione delle performance professionali in contesti operativi reali.

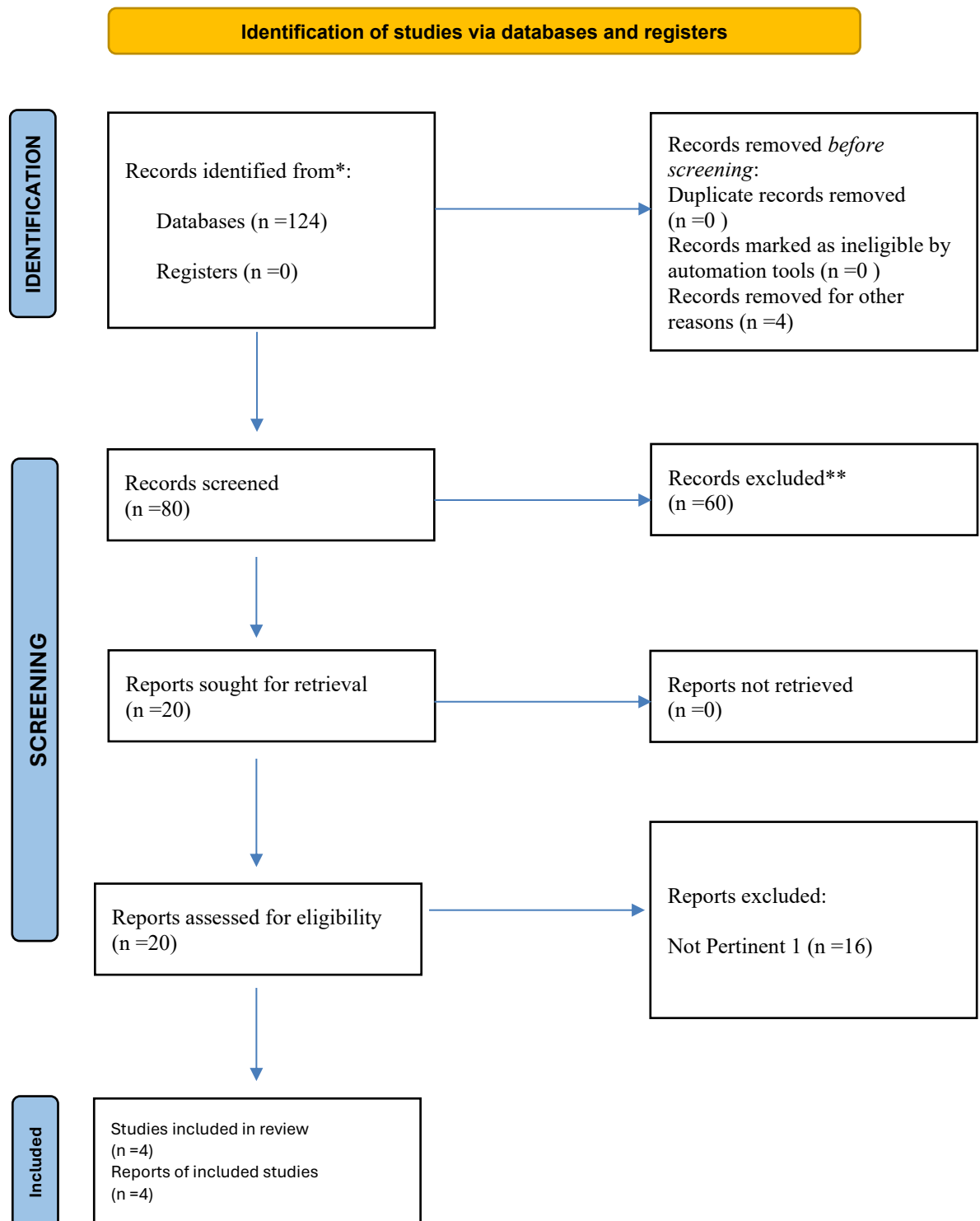
SCOPING REVIEW

A partire dall'individuazione della Piramide di Miller quale principale modello teorico di riferimento, emersa attraverso la fase preliminare di revisione e il successivo processo di snowball search, il percorso metodologico dello studio ha assunto una configurazione maggiormente strutturata e orientata. Tale modello ha rappresentato il riferimento teorico per l'elaborazione di **una nuova stringa di ricerca** (tabella 4), utilizzata per l'elaborazione del diagramma PRISMA [9], orientata all'individuazione di strumenti e sistemi di valutazione basati su approcci competence-based, successivamente utilizzati come base metodologica per la progettazione dello strumento valutativo. Nella seconda fase dell'analisi, l'attenzione si è concentrata sull'esame approfondito del “Sistema di Misurazione e Valutazione della Performance” sviluppato da AGENAS, in conformità all'art. 7, comma 1, del D.Lgs. 150/2009, (aggiornamento 2024) [8] e sulla normativa di riferimento vigente nella Regione Marche. Questo approfondimento è stato finalizzato a

inquadrare il contesto istituzionale e normativo entro cui si colloca l'attività valutativa svolta dai GAAR, e a esplorare la possibilità di integrare modelli teorici di valutazione con i sistemi già esistenti a livello regionale. A seguito dell'applicazione della stringa di ricerca “((*"miller"[All Fields] OR "miller s"[All Fields] OR "millers"[All Fields]*) AND (*"pyramid"[All Fields] OR "pyramidal"[All Fields] OR "pyramidally"[All Fields] OR "pyramidals"[All Fields] OR "pyramids"[All Fields]*)) AND (2015:2025[pdat]) Filters: Free full text, Full text” elaborata sulla base dei termini MeSH, sono stati inizialmente identificati complessivamente **124 articoli** provenienti dai principali database scientifici indicizzati. Nella fase preliminare di identificazione, sono stati esclusi **4 record** per motivi diversi (es. non pertinenza evidente o tipologia di documento non coerente), portando a un totale di **120 articoli potenzialmente eleggibili**. Successivamente, nella fase di screening, sono stati analizzati titoli e abstract di **80 articoli**, mentre **40 articoli** sono stati esclusi prima dello screening per mancata pertinenza o criteri preliminari non soddisfatti. Durante lo screening, sono stati esclusi ulteriori **60 articoli**, in quanto non coerenti con l'obiettivo dello studio, non riferiti alla valutazione delle competenze o non applicabili, anche indirettamente, al contesto sanitario e organizzativo oggetto di indagine. Al termine di questa fase, **20 articoli** sono stati ritenuti idonei per la valutazione full text (reports sought for retrieval). Nella fase di valutazione dell'eleggibilità (full text), sono stati analizzati **20 articoli**, dei quali **16 sono stati esclusi** in quanto non pertinenti rispetto agli obiettivi dello studio, per limitata trasferibilità al contesto sanitario o per insufficiente rilevanza metodologica. Infine, sono stati inclusi nella revisione **4 articoli**, ritenuti metodologicamente solidi e maggiormente pertinenti ai fini della costruzione del framework teorico di riferimento.

Figura 1 Prisma 2020 Diagram

PRISMA 2020 DIAGRAM [9]



*Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers).

**If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.
Source: Page MJ, et al. BMJ 2021;372:n71. doi: 10.1136/bmj.n71.

This work is licensed under CC BY 4.0. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Tabella 4 Tabella Riassuntiva delle fonti bibliografiche

Banca dati	Stringa	Articoli	Selezionati	
pubMED	("miller"[All Fields] OR "miller s"[All Fields] OR "millers"[All Fields]) AND ("pyramid"[All Fields] OR "pyramidal"[All Fields] OR "pyramidally"[All Fields] OR "pyramidals"[All Fields] OR "pyramids"[All Fields])) AND (2015:2025[pdat]) Filters: Free full text, Full text	124	4	1. Cruess, R. L., Cruess, S. R., & Steinert, Y. (2016). Amending Miller's Pyramid to Include Professional Identity Formation. <i>Academic Medicine</i>, 91(2), 180. https://doi.org/10.1097/ACM.0000000000000913 2. Domínguez-Torres, L. C., & Vega-Peña, N. V. (2023). Pyramids in medical education: Their conceptualization and utility summarized. <i>Revista Colombiana De Obstetricia Y Ginecologia</i>, 74(2), 163–174. https://doi.org/10.18597/rcog.3994 3. Miller, G. E. (1990). The assessment of clinical skills/competence/performance. <i>Academic Medicine</i>, 65(9), S63. 4. Rhind, S. M., MacKay, J., Brown, A. J., Mosley, C. J., Ryan, J. M., Hughes, K. J., & Boyd, S. (2021). Developing Miller's Pyramid to Support Students' Assessment Literacy. <i>Journal of Veterinary Medical Education</i>, 48(2), 158–162. https://doi.org/10.3138/jvme.2019-0058

Tabella 5 Summary Of Findings "SOF"

Banca dati: PUBMED					
te	Autori	Obiettivo	Risultati ottenuti	Tipo di studio	Strumenti messi in atto per la valutazione
	Anno pubblicazione				
	Rivista				
Amending Miller's Pyramid to Include Professional Identity Formation	Cruess, Richard L. MD; Cruess, Sylvia R. MD; Steinert, Yvonne PhD	Lo studio ha lo scopo di fornire un quadro strutturato e <i>progressivo</i> (da livello 1 a livello 5, per valutare le competenze cliniche, professionali e comunicative del personale sanitario, sia per la crescita individuale, sia per il miglioramento continuo e per avere un sistema valutativo efficace per valutare le performance raggiunt. Il sistema delle Pediatrics Milestones dell'ACGME si collega perfettamente alla piramide di Miller, che è uno dei modelli più diffusi per descrivere la valutazione delle competenze cliniche nei professionisti sanitari.	Pediatrics Milestones ACGME 2021", non è uno studio sperimentale o osservazionale, ma un documento normativo, valutativo, quindi non presenta risultati sperimentali. È una revisione aggiornata (seconda revisione, 2021) dei criteri utilizzati per valutare la competenza clinica dei medici in formazione pediatrica negli Stati Uniti. Definisce traguardi progressivi che ogni specializzando dovrebbe raggiungere in ciascuna delle 6 competenze fondamentali dell'ACGME. Ogni milestone è suddivisa in 5 livelli, da "principiante" a "esperto", con descrizioni dettagliate delle abilità richieste. Non si lega agli anni di formazione ma al livello reale di performance del valutato, quindi è un sistema		Milestones: framework valutativo strutturato Ogni competenza è suddivisa in sotto competenze. Ogni sotto competenza è descritta attraverso 5 livelli di performance (da principiante a esperto). La valutazione è narrativa e osservazionale, non numerica.
	Febbraio 2016 Rev 2021				
	Academic Medicine Journal of the Association Of American Medical College				

			flessibile, continuo e formativo. Serve come base per le decisioni del Clinical Competency Committee, che monitora lo sviluppo del singolo medico.		Eventuali simulazioni
Developing Miller's Pyramid to Support Students' Assessment Literacy	Susan M. Rhind, Jill MacKay, Andrew J. Brown, Caroline J. Mosley, John M. Ryan, Kirsty J. Hughes, and Sharon	L'obiettivo dello studio è quello di sviluppare e proporre un modello educativo basato sulla piramide di Miller per supportare lo sviluppo della "assessment literacy" negli studenti, ovvero la loro capacità di comprendere, interpretare e utilizzare in modo critico la valutazione durante il percorso formativo. Nello specifico lo studio si propone come obiettivo quello di adattare la Piramide di Miller (Knows → Does) alla formazione degli studenti, non solo per valutare la performance, ma per insegnare come funziona la valutazione stessa.	Lo studio di Rhind et al. (2020), pur non essendo uno studio quantitativo classico, riporta risultati concreti ottenuti attraverso questionari, focus group e osservazioni dopo l'implementazione di attività didattiche mirate allo sviluppo della assessment literacy negli studenti di medicina veterinaria. Aumento della comprensione del processo valutativo Nello specifico, dopo aver partecipato ad attività come la creazione di domande, il PeerWise, la valutazione tra pari e standard setting, 86% degli studenti ha riferito che l'attività ha migliorato la loro comprensione della valutazione.		Standard setting (metodo Angoff) Adaptive Comparative Judgment FOSCE (Formative OSCE)
	Marzo 2020 University Of Toronto.				
Pyramids in medical education: Their conceptualiz	Luis Carlos Domínguez-Torres, Neil Valentín Vega-Peña	Lo studio si propone di analizzare in modo approfondito il significato e l'utilità delle principali piramidi teoriche utilizzate nell'ambito della formazione medica. In particolare, gli autori si concentrano sull'esame dei concetti alla base di questi modelli, sulle critiche che	Lo studio ha portato allo sviluppo di un modello teorico innovativo, la piramide di Miller, che ha rivoluzionato la valutazione delle competenze	Revisione narrativa	"Knows" (conoscenza): Test a scelta multipla (MCQ) Domande a risposta aperta Esami scritti

ation and utility summarized	30 Giugno 2023	hanno ricevuto nel tempo e sulle modifiche che sono state proposte per migliorarli o adattarli alle nuove esigenze formative. L'obiettivo principale è quindi quello di comprendere come queste rappresentazioni grafiche, tra cui spicca la piramide di Miller, possano essere impiegate concretamente per supportare i processi di progettazione curricolare, insegnamento e valutazione nella formazione dei professionisti della salute.	cliniche suddividendole in quattro livelli progressivi: <i>knows, knows how, shows how e does</i> . Questo modello ha permesso di comprendere che nessun singolo metodo può valutare da solo la complessità della pratica medica.		"Knows how" (competenza applicativa): Analisi di casi clinici Prove di problem solving Essay test "Showhow" (dimostrazione in ambiente simulato): OSCE(Objective Structured Clinical Examination) Pazienti standardizzati Simulazioni pratiche "Does" (performance reale): Valutazioni sul campo (workplace-based assessment) Mini-CEX Feedback 360° Audit clinici Supervisione diretta Questi strumenti sono impiegati in modo integrato per garantire una valutazione completa delle competenze professionali in ambito medico, sia teoriche che pratiche, individuali e relazionali.
	Revista Colombiana de Obstetricia y Ginecología Vol. 74 No.2 - Abril-Junio 2023 - (163-174)	Lo studio intende anche sottolineare l'importanza di valutare l'impatto reale di tali modelli nei contesti educativi locali, ponendo l'attenzione sulla necessità di una maggiore consapevolezza da parte dei docenti rispetto all'evidenza scientifica che li sostiene.	Tra i risultati principali: Ha evidenziato i limiti dei test teorici, sostenendo la necessità di strumenti più validi per misurare le capacità pratiche e comportamentali.		
		Lo studio ha analizzato cinque piramidi educative (Miller, Bloom, Kirkpatrick, Dale e Maslow), evidenziandone i fondamenti teorici, le applicazioni nella formazione medica e le principali critiche. È emerso che, nonostante alcune limitazioni, questi modelli risultano ancora utili per guidare la progettazione curricolare, la didattica e la valutazione. In particolare, la piramide di Miller ha stimolato riflessioni sull'identità professionale e sull'affidabilità clinica, portando a proposte di ampliamento del modello originario. Gli autori sottolineano infine l'importanza di valutarne l'efficacia nei contesti locali e di promuoverne un uso più consapevole da parte dei formatori.	Ha promosso l'uso di strumenti innovativi come pazienti standardizzati e OSCE, dimostrando che possono offrire una valutazione più affidabile e realistica delle competenze cliniche. Ha contribuito a orientare l'educazione medica verso un approccio basato sulle competenze (competency-based). Ha stimolato l'evoluzione di sistemi di valutazione più integrati, criteriali e formativi, influenzando pratiche educative a livello internazionale. In sintesi, il lavoro di Miller ha offerto una struttura concettuale chiara, ancora oggi alla base della		

			valutazione clinica moderna, aprendo la strada a studi successivi e alla revisione dei curricula di medicina.		
The assessment of clinical skills/competence/performance	G E Miller Settembre 1990	L'obiettivo principale di questo studio è quello di proporre un modello concettuale articolato per la valutazione delle competenze cliniche, attraverso un approccio piramidale che distingue e collega progressivamente i diversi livelli della professionalità medica: dalla semplice conoscenza teorica ("knows"), alla capacità di applicarla ("knows how"), fino alla dimostrazione pratica delle competenze in contesti simulati ("shows how") e, infine, alla performance reale nel contesto professionale ("does"). A partire da questa cornice teorica, lo studio intende analizzare in modo critico i principali strumenti e metodi utilizzati per valutare ciascun livello di competenza, evidenziandone punti di forza e debolezza, nonché il grado di affidabilità, validità e rilevanza per la pratica clinica. In particolare, vengono esaminati metodi oggettivi per la misurazione della conoscenza (test a scelta multipla), strumenti per valutare l'applicazione del ragionamento clinico (come le domande a risposta strutturata, i MEQ e i PMP), simulazioni e pazienti standardizzati per valutare la performance simulata, fino all'osservazione diretta nei contesti clinici reali. Lo studio mira inoltre a superare la visione tradizionale secondo cui la valutazione della preparazione medica si possa esaurire con la sola misurazione del sapere teorico. Al contrario, viene fortemente sottolineata la necessità di integrare, nei percorsi di formazione e accreditamento, strumenti in grado di misurare ciò che il professionista sa	Lo studio ha portato alla definizione di un modello piramidale per la valutazione delle competenze cliniche, distinguendo tra conoscenza, competenza, performance simulata e azione reale. È emerso che i tradizionali test teorici, pur affidabili, sono insufficienti per valutare le reali capacità cliniche. L'introduzione di strumenti come i pazienti standardizzati e l'OSCE ha dimostrato maggiore efficacia nella valutazione della performance, con buona validità e accettabile affidabilità, soprattutto se accompagnati da checklist standardizzate e valutatori formati. Si è rilevato che anche osservatori non-medici possono contribuire in modo valido alla valutazione. È stata evidenziata la necessità di superare l'approccio normativo, promuovendo valutazioni criteriali orientate	Revisione narrativa e concettuale	Miller ha adottato una strategia di revisione concettuale basata sull'analisi critica della letteratura e sull'esperienza professionale. Ha elaborato un modello teorico (la piramide di Miller) per classificare i livelli di competenza clinica e ha confrontato diversi metodi valutativi, evidenziandone affidabilità, validità e limiti. La strategia include anche una riflessione sulle implicazioni educative e sull'uso integrato di strumenti innovativi come i pazienti standardizzati e l'OSCE.
	Academic Medicine. Volume 65. N.9				

		effettivamente fare, e non solo ciò che conosce. In quest'ottica, viene sostenuta con forza l'adozione di strumenti valutativi più realistici, standardizzati e affidabili, come l'uso dei pazienti simulati e delle OSCE (Objective Structured Clinical Examinations), capaci di riprodurre fedelmente situazioni cliniche reali e di fornire una valutazione più completa delle capacità comunicative, tecniche e decisionali del candidato. Un ulteriore obiettivo dello studio è quello di promuovere un cambiamento culturale nel sistema di valutazione, suggerendo il passaggio da un modello normativo – che classifica gli studenti in base al rendimento relativo – a un modello criterioale, volto a verificare il raggiungimento di standard minimi e imprescindibili di competenza professionale, più coerente con la responsabilità sociale di chi forma i futuri medici. Infine, lo studio sottolinea l'importanza della formazione dei valutatori, della standardizzazione degli strumenti di osservazione (checklist e scale di valutazione), e della necessità di un impegno collettivo, da parte della comunità accademica, per sviluppare sistemi di valutazione più efficaci, trasparenti e sostenibili, specialmente in vista dell'applicazione in contesti ad alto rischio decisionale, come gli esami di abilitazione o le procedure di accreditamento professionale.	a certificare il raggiungimento di competenze minime attese. Infine, lo studio suggerisce l'adozione di strategie valutative più sostenibili, come il testing sequenziale, e invita a proseguire la ricerca per migliorare l'efficacia e l'equità dei metodi di valutazione clinica.		
--	--	---	---	--	--

2° FASE: VALIDAZIONE DELLA SCHEDA DI VALUTAZIONE

FOCUS GROUP PER LA COSTRUZIONE E VALIDAZIONE PRELIMINARE DELLA SCHEDA DI VALUTAZIONE

Nell'ambito della seconda fase dello studio, orientata alla progettazione applicativa e alla validazione preliminare dello strumento valutativo, il focus group composto da 4 membri dell'OTA della Regione MARCHE organizzato il 27 Novembre 2025 ha rappresentato il primo momento metodologico strutturato di confronto finalizzato alla traduzione operativa delle evidenze teoriche emerse dalla revisione della letteratura in un sistema di valutazione concretamente applicabile al contesto dei valutatori del Gruppo di Autorizzazione e Accreditamento Regionale (GAAR). Tale fase si è resa necessaria in considerazione della natura prevalentemente metodologica del progetto, il cui obiettivo non era esclusivamente individuare modelli teorici di riferimento, ma sviluppare uno strumento valutativo contestualizzato, coerente con le specificità operative, normative e organizzative dell'Organismo Tecnicamente Accreditante (OTA) della Regione Marche. Il focus group è stato concepito come metodologia qualitativa di confronto tra esperti, finalizzata a integrare il sapere derivante dalla letteratura scientifica con l'esperienza professionale maturata sul campo dai soggetti direttamente coinvolti nei processi di autorizzazione, accreditamento e valutazione. In termini metodologici, questa scelta risponde alla necessità di garantire non soltanto validità teorica allo strumento, ma anche coerenza e applicabilità pratica, elementi imprescindibili nella costruzione di sistemi valutativi destinati a contesti organizzativi complessi. La letteratura, infatti, evidenzia come i processi di sviluppo di strumenti di valutazione delle competenze risultino maggiormente efficaci quando fondati su una combinazione tra evidenze scientifiche, framework teorici validati e contributo esperienziale degli stakeholder coinvolti nel processo applicativo. Il focus group è stato pertanto strutturato come momento di confronto tecnico-scientifico multidisciplinare, finalizzato ad analizzare criticamente la scheda di valutazione precedentemente in uso, identificandone punti di forza, limiti metodologici e aree di possibile sviluppo. In particolare, il confronto si è concentrato sulla capacità dello strumento esistente di rappresentare in maniera sufficientemente articolata la complessità della performance del valutatore GAAR, considerando non solo l'adempimento delle attività assegnate, ma anche dimensioni più ampie quali la qualità

metodologica, la capacità applicativa, la partecipazione ai processi istituzionali, la collaborazione organizzativa e il contributo al miglioramento complessivo del sistema. L'analisi condivisa ha evidenziato come il sistema precedentemente adottato, pur garantendo una base strutturata di valutazione, risultasse prevalentemente orientato alla verifica dell'esecuzione delle attività e meno focalizzato sulla progressione delle competenze, sulla multidimensionalità della performance e sulla piena esplicitazione del razionale teorico sottostante ai criteri di giudizio. Tale riflessione ha orientato il gruppo verso la necessità di sviluppare uno strumento maggiormente strutturato, capace di integrare la dimensione teorica competence-based, rappresentata dalla Piramide di Miller, con la dimensione organizzativa e prestazionale derivante dal Sistema di Misurazione e Valutazione della Performance di AGENAS. Dal punto di vista operativo, il focus group ha consentito di definire i principali assi metodologici della nuova scheda. In primo luogo, sono state identificate le macroaree valutative ritenute maggiormente rappresentative delle funzioni esercitate dal valutatore GAAR, successivamente organizzate secondo una logica coerente con le diverse fasi del processo di audit e con le responsabilità istituzionali connesse al ruolo. In secondo luogo, è stata avviata la selezione degli item valutativi, orientata alla costruzione di indicatori osservabili, misurabili e chiaramente associabili a specifiche dimensioni di performance. Questo passaggio ha rappresentato un elemento centrale del processo, in quanto ha consentito di tradurre costrutti teorici complessi in componenti valutative concretamente applicabili. Un ulteriore aspetto di rilievo ha riguardato la definizione delle scale di misurazione. Attraverso il confronto tra i partecipanti, è emersa la necessità di adottare un sistema di scoring capace di ridurre l'ambiguità interpretativa, favorire la standardizzazione del giudizio e incrementare la riproducibilità delle valutazioni. La riflessione metodologica si è pertanto orientata verso la costruzione di livelli di valutazione chiaramente descritti, associati a differenti gradi di performance e progressione delle competenze, in coerenza con il framework teorico adottato. Il focus group ha inoltre svolto una funzione di validazione preliminare del contenuto, consentendo di verificare la pertinenza degli item proposti rispetto alle reali attività svolte dai GAAR e al contesto normativo-organizzativo regionale. Tale passaggio è risultato fondamentale per evitare il rischio di costruire uno strumento teoricamente robusto ma scarsamente aderente alla pratica operativa. In questa prospettiva, il confronto tra dimensione teorica e dimensione esperienziale ha rappresentato uno dei principali

punti di forza dell'intero processo. Complessivamente, il focus group si configura come il passaggio metodologico che ha consentito la trasformazione del framework teorico elaborato nella prima fase dello studio in una struttura valutativa concretamente applicabile, fungendo da ponte tra revisione della letteratura e implementazione sperimentale. Questo processo ha consentito di consolidare la solidità metodologica e scientifica del progetto, rafforzandone la trasferibilità al contesto operativo e ponendo le basi per le successive fasi di sperimentazione e validazione sul campo. In tale prospettiva, il focus group non ha rappresentato un semplice momento consultivo, bensì una fase strategica di co-costruzione metodologica dello strumento, finalizzata a garantire che la nuova scheda di valutazione fosse non solo teoricamente fondata, ma anche condivisa, applicabile e coerente con le esigenze istituzionali dell'OTA e con la complessità professionale dei valutatori GAAR.

LA SCHEDA DI VALUTAZIONE “MILLER - AGENAS”

Sulla base delle evidenze emerse dalla revisione della letteratura e dall'analisi dei modelli teorici e istituzionali individuati (MILLER-AGENAS), il risultato finale del focus group è stato lo sviluppo di uno strumento strutturato per la valutazione delle competenze dei valutatori del Gruppo di Autorizzazione e Accreditamento Regionale (GAAR). Tale strumento rappresenta l'esito del percorso metodologico descritto nelle fasi precedenti e si configura come una sintesi operativa dei riferimenti teorici e degli elementi emersi dall'analisi critica del contesto valutativo. La scheda (**allegato 3**) è stata concepita come uno strumento applicabile nella pratica operativa, caratterizzato da una struttura organizzata e da criteri di valutazione esplicitati, al fine di consentire una rilevazione sistematica e standardizzata delle performance. La sua elaborazione ha seguito un approccio coerente con i modelli di valutazione delle competenze e con i sistemi di misurazione della performance analizzati, garantendo un allineamento tra dimensione teorica e applicazione operativa. Pertanto la scheda di valutazione sviluppata nell'ambito del presente studio si configura come uno strumento strutturato, analitico e multidimensionale. La sua progettazione risponde all'esigenza di superare approcci valutativi prevalentemente descrittivi o non pienamente standardizzati, introducendo un modello fondato su criteri osservabili, indicatori espliciti e livelli di giudizio progressivi, in grado di garantire maggiore oggettività, coerenza metodologica e riproducibilità del

processo valutativo. Dal punto di vista strutturale, la scheda è articolata in quattro aree valutative principali (A, B, C e D), ciascuna delle quali corrisponde a una specifica fase del processo di autorizzazione e accreditamento e consente di analizzare in maniera distinta ma integrata le diverse dimensioni della performance professionale del valutatore. Tale articolazione non è casuale, ma riflette una precisa scelta metodologica, orientata a ricostruire l'intero ciclo operativo dell'attività valutativa, dalla fase preparatoria fino alla dimensione organizzativa e partecipativa, permettendo così una valutazione globale e contestualizzata.

L'Area A – Performance operativa in fase di pre-audit rappresenta il primo momento del processo valutativo ed è finalizzata ad analizzare le competenze espresse nella fase preparatoria. In questa sezione vengono valutati aspetti fondamentali quali la capacità di pianificare il pre-audit nel rispetto delle tempistiche previste, l'accuratezza nella redazione dei verbali preliminari e la qualità dell'analisi delle evidenze documentali fornite dalla struttura. Tali elementi assumono un ruolo determinante in quanto la fase di pre-audit costituisce il presupposto metodologico dell'intero processo: una pianificazione inadeguata o un'analisi superficiale delle evidenze può compromettere la qualità delle fasi successive, incidendo negativamente sulla coerenza e sull'efficacia dell'audit.

L'Area B – Performance operativa in fase di audit si configura come il nucleo centrale della valutazione, in quanto prende in esame la capacità del valutatore di operare direttamente sul campo durante la visita ispettiva. In questa sezione vengono analizzati la qualità e la completezza dei verbali di audit, la capacità di esaminare in maniera sistematica e metodologicamente corretta le evidenze raccolte e, soprattutto, la coerenza tra le evidenze osservate e i giudizi formulati. Quest'ultimo aspetto rappresenta un indicatore particolarmente rilevante di competenza professionale avanzata, poiché implica non solo la capacità di osservare e registrare dati, ma anche quella di interpretarli criticamente e tradurli in valutazioni tecniche fondate, evitando discrepanze, sovrastime o sottovalutazioni. In tal senso, l'Area B consente di cogliere il passaggio da una conoscenza teorica a una competenza applicata e contestualizzata.

L'Area C – Conclusione della visita di audit è dedicata alla fase finale del processo valutativo e si concentra sulla formalizzazione degli esiti dell'attività ispettiva. In questa

sezione vengono valutati la completezza e la coerenza delle prescrizioni formulate, la correttezza nella compilazione della documentazione amministrativa e il rispetto delle tempistiche di consegna dei verbali sia alla struttura valutata sia all'Organismo Tecnicamente Accreditante. Questa area assume una particolare rilevanza in termini di responsabilità professionale e accountability, in quanto la qualità della documentazione prodotta rappresenta non solo un output tecnico, ma anche un atto formale con implicazioni organizzative, amministrative e legali. Una documentazione incompleta o tardiva può infatti compromettere la trasparenza e la credibilità del processo di accreditamento.

L'Area D – Partecipazione e collaborazione istituzionale amplia ulteriormente la prospettiva valutativa, includendo dimensioni trasversali legate al comportamento organizzativo e al contributo del valutatore al funzionamento complessivo del sistema. In questa sezione vengono analizzati elementi quali la partecipazione agli eventi formativi, il coinvolgimento nelle riunioni dell'OTA e la disponibilità a partecipare alle attività ispettive. Tale area consente di valutare aspetti quali l'impegno professionale, la continuità della partecipazione e il grado di integrazione del valutatore all'interno dell'organizzazione, riconoscendo che la qualità del sistema non dipende esclusivamente dalle prestazioni individuali durante l'audit, ma anche dalla partecipazione attiva ai processi di miglioramento continuo e aggiornamento professionale.

Dal punto di vista della misurazione, ciascun item è valutato mediante una scala Likert a quattro livelli (2, 4, 6, 8), scelta in modo deliberato per evitare la presenza di un punto centrale neutro e favorire una maggiore capacità discriminante del giudizio. L'assenza di un valore intermedio obbliga infatti il valutatore a esprimere un orientamento più definito rispetto allo standard atteso, riducendo il rischio di valutazioni ambigue o eccessivamente prudenti. In questo contesto, il punteggio 2 indica una prestazione significativamente insufficiente rispetto agli standard richiesti, il punteggio 4 segnala un livello parzialmente adeguato ma con criticità rilevanti, il punteggio 6 rappresenta una prestazione complessivamente adeguata ma migliorabile, mentre il punteggio 8 identifica il pieno raggiungimento degli standard, con elevati livelli di accuratezza, coerenza e qualità metodologica. I punteggi attribuiti ai singoli item vengono aggregati all'interno delle rispettive aree valutative, ciascuna delle quali presenta un punteggio massimo pari a 24.

Tale scelta consente di attribuire pari peso alle diverse fasi del processo valutativo, evitando squilibri tra le dimensioni analizzate e garantendo una valutazione complessiva equilibrata. Il punteggio totale, risultante dalla somma delle quattro aree (range 0–96), consente di collocare il valutatore all'interno di specifici livelli di giudizio complessivi. La classificazione finale è articolata in quattro livelli: “Non idoneo” (0–24), “Parzialmente idoneo” (25–48), “Idoneo” (49–72) ed “Eccellente” (73–96). Tali livelli sono stati definiti attraverso un'integrazione metodologica tra il Sistema di Misurazione e Valutazione della Performance e la Piramide di Miller, consentendo di associare il punteggio numerico a un livello progressivo di competenza. In particolare, il livello “Parzialmente idoneo” è riconducibile al livello “Knows” della Piramide di Miller, indicativo di una conoscenza teorica ancora non pienamente tradotta in competenza operativa; il livello “Idoneo” si colloca tra i livelli “Knows How” e “Shows How”, riflettendo una capacità applicativa e dimostrativa della competenza; il livello “Eccellente”, infine, corrisponde al livello “Does”, che rappresenta la piena espressione della competenza nella pratica reale, caratterizzata da autonomia, coerenza e capacità di integrazione delle diverse dimensioni della performance.

VALIDAZIONE PRELIMINARE DELLA NUOVA SCHEDA DI VALUTAZIONE “MILLER AGENAS”.

A conclusione del processo di elaborazione metodologica sviluppato attraverso il focus group, la nuova scheda di valutazione delle performance è stata formalmente condivisa con i soggetti istituzionalmente coinvolti nel sistema di governance dell'Organismo Tecnicamente Accreditante, al fine di verificarne coerenza metodologica, aderenza al contesto operativo e compatibilità con le procedure regionali vigenti. Tale fase di condivisione ha rappresentato un passaggio fondamentale nel percorso di sviluppo dello strumento, consentendo una verifica finale della struttura elaborata, degli item selezionati, delle scale di misurazione adottate e della loro applicabilità concreta all'interno del processo di valutazione dei GAAR. A seguito di tale confronto, la scheda è stata sottoposta all'approvazione del Coordinatore dell'OTA, quale figura responsabile della supervisione tecnica e metodologica del sistema valutativo regionale. Questa fase di validazione preliminare è stata concepita non come un mero adempimento formale, bensì come un processo metodologico finalizzato a verificare la solidità concettuale dello

strumento, la pertinenza delle dimensioni valutative individuate e la coerenza tra il framework teorico di riferimento e la sua traduzione operativa. Tale passaggio ha consentito di sottoporre la scheda a una verifica critica preliminare, orientata ad accertarne la chiarezza, la completezza e l'effettiva applicabilità nel contesto delle attività di valutazione svolte dai GAAR, prima della successiva implementazione sul campo. L'approvazione ha sancito la validazione istituzionale preliminare dello strumento, autorizzandone l'introduzione nel contesto operativo e la successiva fase di sperimentazione applicativa. Questo passaggio ha conferito alla nuova scheda non solo legittimazione formale, ma anche piena integrazione nel sistema organizzativo di riferimento, rappresentando il punto di transizione tra la fase progettuale e la successiva implementazione sul campo.

IMPLEMENTAZIONE OPERATIVA DELLA SCHEDA DI VALUTAZIONE “MILLER AGENAS”.

A seguito della conclusione del processo di progettazione metodologica, della condivisione tecnico-istituzionale e della validazione preliminare da parte del Coordinatore dell'Organismo Tecnicamente Accreditante (OTA), la scheda di valutazione “Miller–AGENAS” è stata formalmente introdotta nella pratica operativa quale strumento sperimentale per la valutazione annuale delle performance dei valutatori del Gruppo di Autorizzazione e Accreditamento Regionale (GAAR) della Regione Marche. L'implementazione dello strumento ha rappresentato il passaggio applicativo attraverso cui il modello teorico-metodologico sviluppato nelle fasi precedenti è stato trasferito dal piano progettuale al contesto istituzionale reale, consentendone la sperimentazione sul campo in un setting organizzativo concreto. Nello specifico, la scheda è stata applicata all'intera popolazione dei valutatori GAAR attivi iscritti nell'elenco regionale della Regione Marche per l'anno 2025, corrispondente a un totale di 115 professionisti. Tale scelta metodologica ha consentito di estendere l'utilizzo dello strumento all'intero universo di riferimento, evitando processi di campionamento e garantendo una valutazione sistematica, omogenea e contestualizzata dell'intera rete regionale dei valutatori. L'applicazione su scala completa ha inoltre permesso di verificare la sostenibilità operativa, la praticabilità organizzativa e la capacità dello strumento di adattarsi alle differenti caratteristiche professionali dei soggetti valutati,

mantenendo uniformità metodologica e coerenza valutativa. L'implementazione operativa è stata condotta nell'ambito del consueto processo di valutazione annuale previsto per il mantenimento delle competenze dei valutatori, configurandosi pertanto non come attività separata o esclusivamente sperimentale, ma come integrazione strutturata di un nuovo modello valutativo all'interno di una procedura istituzionale già formalizzata. In questa prospettiva, la scheda "Miller–AGENAS" ha sostituito, in via sperimentale, il precedente strumento di valutazione, introducendo un sistema maggiormente articolato, multidimensionale e fondato su criteri osservabili, indicatori strutturati e livelli di performance coerenti con i framework teorici adottati.

L'utilizzo della nuova scheda ha consentito di valutare in maniera più sistematica non soltanto l'adempimento delle attività assegnate, ma anche la progressione delle competenze, la qualità metodologica, la partecipazione ai processi istituzionali, la collaborazione organizzativa e la coerenza professionale rispetto agli standard richiesti dal ruolo. In tal senso, la fase di implementazione ha costituito non solo un momento di applicazione tecnica dello strumento, ma anche una verifica preliminare della sua capacità di operare come dispositivo di misurazione delle performance all'interno di un sistema regionale complesso. Dal punto di vista metodologico, l'applicazione all'intera popolazione GAAR ha rappresentato una fase cruciale del percorso di validazione sul campo, consentendo di osservare il comportamento dello strumento in condizioni operative reali e di raccogliere evidenze preliminari in merito alla sua applicabilità, coerenza e potenziale trasferibilità. L'implementazione della scheda "Miller–AGENAS" nel processo valutativo 2025 si configura pertanto come il primo concreto utilizzo istituzionale del modello sviluppato, costituendo il fondamento empirico per le successive fasi di raccolta dati, analisi dei risultati e valutazione del gradimento dei professionisti coinvolti.

QUESTIONARIO PER LA VALUTAZIONE PERCETTIVA DELLA SCHEDA DI VALUTAZIONE "MILLER AGENAS"

Al fine di accompagnare la fase di implementazione operativa della scheda "Miller–AGENAS" con una valutazione preliminare della sua percezione da parte dei soggetti direttamente coinvolti nel processo applicativo, è stato progettato e somministrato il 31 Dicembre 2025 contestualmente all'invio delle valutazioni individuali a ciascun

valutatore GAAR, al fine di garantire trasparenza del processo valutativo, un questionario di gradimento rivolto ai valutatori GAAR sottoposti al nuovo sistema valutativo. La costruzione di tale strumento è stata finalizzata a integrare la dimensione oggettiva derivante dall'applicazione sperimentale della scheda con una componente percettiva qualitativo-quantitativa, orientata a rilevare il grado di chiarezza, comprensibilità, utilità e accettabilità del modello proposto. In questa prospettiva, il questionario ha rappresentato un dispositivo metodologico complementare, concepito non come semplice rilevazione di soddisfazione, ma come strumento di raccolta strutturata di feedback utile a esplorare la sostenibilità applicativa del nuovo sistema, la coerenza percepita rispetto alle attività valutative svolte e l'eventuale presenza di aree suscettibili di ulteriore miglioramento. La sua elaborazione si inserisce pertanto nel più ampio processo di validazione preliminare dello strumento, contribuendo a rafforzare la dimensione partecipativa del progetto e a fornire elementi empirici utili per le successive fasi di analisi e sviluppo. Il questionario aveva l'obiettivo di raccogliere feedback qualitativi e quantitativi in merito alla chiarezza degli item di valutazione, alla comprensibilità delle scale di misurazione, alla percezione di equità e trasparenza del nuovo sistema e alla sua effettiva applicabilità nel contesto operativo, esso è composto da sette item complessivi. I primi cinque quesiti sono strutturati secondo una scala Likert a cinque livelli di risposta:

1 = per nulla;

2 = poco;

3 = abbastanza;

4 = molto;

5 = completamente

e sono finalizzati a indagare diversi aspetti del nuovo modello valutativo, tra cui la chiarezza complessiva del sistema, l'utilità dell'articolazione della valutazione nelle diverse fasi operative dell'attività GAAR (pre-audit, audit e conclusione della visita), la comprensibilità dei criteri di valutazione, la pertinenza dell'inclusione della dimensione di partecipazione e collaborazione istituzionale e il contributo dello strumento al miglioramento professionale del valutatore. Gli ultimi due quesiti prevedono invece una risposta dicotomica (sì/no) e sono finalizzati a rilevare il livello di familiarità pregressa dei valutatori con due modelli di riferimento utilizzati nella costruzione dello strumento:

la Piramide di Miller, quale framework per la valutazione delle competenze professionali, e il sistema di valutazione promosso da AGENAS. Infine, il questionario include una domanda aperta conclusiva, finalizzata a raccogliere eventuali suggerimenti o osservazioni utili al miglioramento del modello di valutazione. Tale sezione qualitativa consente di integrare i dati quantitativi con contributi descrittivi provenienti dall'esperienza diretta dei valutatori coinvolti nel processo (**allegato 4**). Nello specifico li quesiti del questionario riguardano:

- la **chiarezza complessiva del nuovo sistema di valutazione delle competenze;**

attraverso la domanda: *“Nel complesso, quanto ritieni chiaro il nuovo sistema di valutazione delle competenze?”*;

- l'**utilità dell'articolazione del processo valutativo nelle diverse fasi operative (pre-audit, audit e conclusione della visita);**

mediante la domanda: *“Ritieni utile articolare la valutazione nelle diverse fasi dell'attività GAAR (pre-audit, audit, conclusione della visita)?”*;

- la **comprensibilità dei criteri di valutazione adottati;**

attraverso la domanda: *“Quanto ritieni chiari i criteri di valutazione utilizzati nel nuovo modello di valutazione delle performance GAAR?”*;

- la **pertinenza dell'inclusione della dimensione di partecipazione e collaborazione istituzionale;**

mediante la domanda: *“Quanto ritieni pertinente includere la partecipazione e collaborazione istituzionale tra i criteri di valutazione delle performance GAAR?”*;

- il **contributo dello strumento al miglioramento professionale del valutatore;**

attraverso la domanda: *“Quanto ritieni pertinente che il nuovo modello di valutazione delle performance sia utile per favorire il miglioramento professionale del valutatore?”*.

Gli ultimi due quesiti invece come detto in precedenza sono stati strutturati con risposta dicotomica (sì/no) e sono finalizzati a rilevare il livello di familiarità pregressa dei valutatori con i modelli di riferimento utilizzati nella costruzione dello strumento. In particolare, le domande hanno riguardato:

- la conoscenza della **Piramide di Miller**, quale modello teorico per la valutazione delle competenze professionali;
- la conoscenza del **Sistema di Misurazione e Valutazione della Performance promosso da AGENAS**.

La raccolta di tali informazioni ha consentito di integrare i dati derivanti dall'applicazione dello strumento con le percezioni dei soggetti valutati, fornendo ulteriori elementi utili per una valutazione preliminare dell'efficacia e dell'accettabilità della scheda, nonché per l'eventuale individuazione di margini di miglioramento in vista della sua implementazione stabile nei processi istituzionali di valutazione dei valutatori GAAR.

METODOLOGIA DI ANALISI DEI DATI QUALITATIVI OTTENUTI

I risultati del questionario di gradimento sono stati analizzati in maniera preliminare mediante elaborazione delle frequenze di risposta per ciascun item e rappresentati graficamente, al fine di consentire una lettura immediata della distribuzione delle risposte prima dell'effettiva analisi statistica dei risultati stessi. Per ciascun quesito sono stati costruiti grafici descrittivi che riportano la distribuzione percentuale delle risposte secondo la scala Likert a cinque livelli, permettendo di evidenziare le tendenze prevalenti e il grado di consenso espresso dai valutatori. Per quanto riguarda le risposte fornite alla domanda aperta, esse sono state oggetto di analisi qualitativa mediante un processo di categorizzazione tematica, finalizzato all'individuazione dei principali nuclei di significato emergenti dalle osservazioni dei partecipanti. Le risposte fornite alla domanda aperta sono state sottoposte a content analysis descrittiva mediante codifica tematica non mutuamente esclusiva, che ha consentito di individuare sei principali categorie di contenuto. La prima categoria, rappresentata dall'assenza di suggerimenti o dalla conferma del modello, comprende le risposte nelle quali i partecipanti non hanno evidenziato proposte di modifica o hanno espresso una sostanziale approvazione dello

strumento sviluppato. La seconda categoria riguarda la valutazione positiva e l'adeguatezza percepita del nuovo sistema di valutazione, includendo osservazioni orientate a sottolinearne la completezza, la coerenza e la pertinenza rispetto alle attività dei valutatori GAAR. Una terza categoria è costituita dai suggerimenti relativi all'affinamento dei criteri di valutazione, della metodologia e degli indicatori, finalizzati a incrementare ulteriormente la chiarezza e l'efficacia dello strumento. La quarta categoria comprende le osservazioni riconducibili al tema del confronto, del feedback e della partecipazione del valutato, evidenziando l'importanza di promuovere momenti di condivisione e restituzione nell'ambito del processo valutativo. La quinta categoria raccoglie i contributi riferiti al monitoraggio longitudinale, alla formazione e al miglioramento continuo, con particolare attenzione alla necessità di prevedere un aggiornamento periodico dello strumento e percorsi di sviluppo professionale coerenti con la sua applicazione. Infine, una sesta categoria, numericamente residuale, riguarda il benchmarking esterno, comprendendo proposte orientate al confronto con modelli e pratiche valutative adottati in altri contesti organizzativi e istituzionali. **(allegato 5).**

METODI STATISTICI

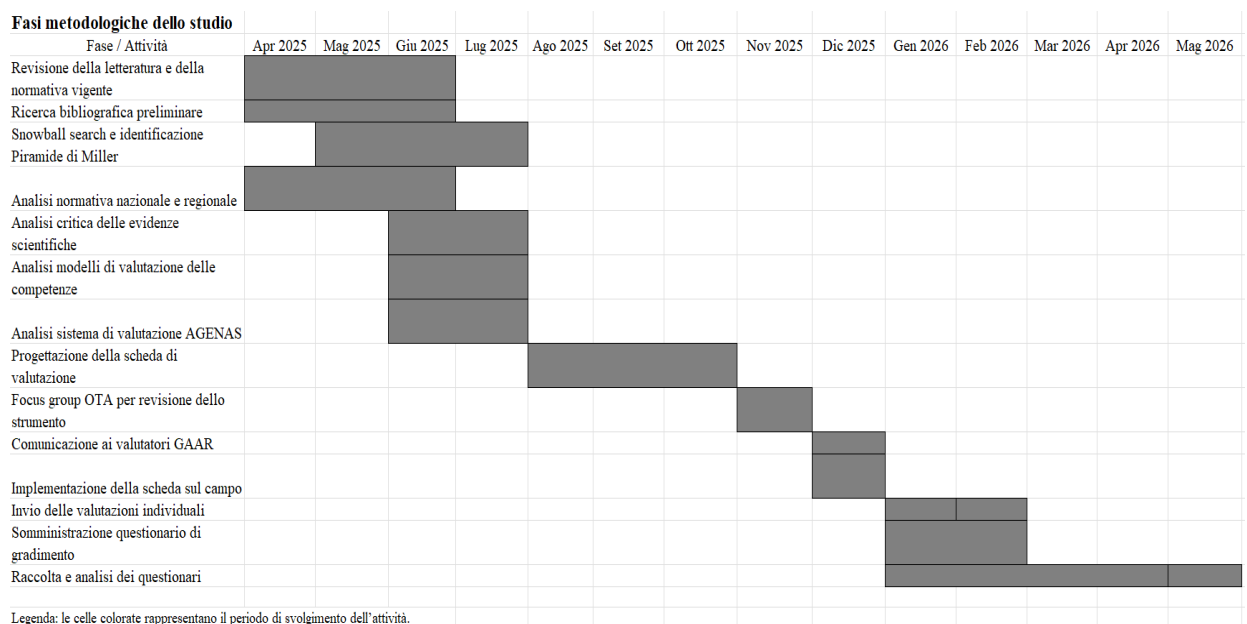
Le variabili Q1–Q5 sono state trattate come item ordinali su scala Likert a cinque livelli (1=per nulla; 2=poco; 3=abbastanza; 4=molto; 5=completamente). È stato costruito uno score composito di percezione/accettabilità del nuovo sistema di valutazione, calcolato come media aritmetica degli item Q1–Q5. Le variabili Q6 e Q7, relative alla familiarità con la Piramide di Miller e con il sistema AGENAS, sono state codificate ordinalmente (0=No; 1=In parte; 2=Si).

L'analisi descrittiva ha previsto frequenze assolute e percentuali per le variabili categoriche e media, deviazione standard, mediana, range interquartile e intervallo di confidenza al 95% per gli item numerici. La normalità dello score composito è stata verificata mediante test di Shapiro-Wilk. Considerata la natura ordinale delle variabili e la distribuzione non normale dello score, le analisi inferenziali bivariate sono state condotte mediante test non parametrici: Wilcoxon signed-rank test per il confronto con il valore neutro della scala (3=abbastanza), Kruskal-Wallis per il confronto dello score tra

livelli di familiarità e coefficiente rho di Spearman per le correlazioni tra item. L'associazione tra familiarità con Piramide di Miller e familiarità con sistema AGENAS è stata valutata mediante test χ^2 (test del chi2).

La consistenza interna dei cinque item è stata valutata mediante α di Cronbach e correlazioni item-totale corrette. La dimensionalità dello strumento è stata esplorata mediante analisi fattoriale esplorativa su matrice di correlazione, preceduta da indice KMO e test di sfericità di Bartlett. È stata infine stimata una regressione lineare multipla esplorativa, con score composito come variabile dipendente e familiarità con Miller e AGENAS come predittori ordinali. Data la non-normalità dei residui e la distribuzione asimmetrica dello score, sono stati riportati errori standard robusti HC3. Il livello di significatività statistica è stato fissato a $p < 0,05$.

Figura 2 Cronoprogramma delle fasi metodologiche



RISULTATI

DISTRIBUZIONE DELLE RISPOSTE AI QUESITI DEL QUESTIONARIO DI GRADIMENTO

Nel corso della fase sperimentale, la nuova scheda di valutazione “Miller–AGENAS” è stata applicata all’intera popolazione dei valutatori GAAR attivi della Regione Marche per l’anno 2025, corrispondente a 115 soggetti. Contestualmente alla trasmissione delle valutazioni individuali, è stato somministrato un questionario di gradimento finalizzato a raccogliere elementi preliminari relativi alla percezione, alla chiarezza e all’applicabilità del nuovo sistema valutativo. Complessivamente sono stati restituiti 73 questionari compilati, corrispondenti al 63,5% della popolazione eleggibile, dato che evidenzia un buon livello di adesione e partecipazione al processo di valutazione sperimentale. Di seguito vengono presentati i risultati preliminari dell’analisi descrittiva dei singoli quesiti del questionario di gradimento, mediante rappresentazione grafica della distribuzione delle risposte fornite dai valutatori GAAR. Tale presentazione consente di evidenziare in modo immediato e preliminare l’orientamento complessivo dei partecipanti rispetto ai principali domini indagati, offrendo una lettura sintetica della percezione del nuovo sistema valutativo, che successivamente sarà sottoposto ad una analisi statistica dettagliata.

Figura 3 Grafico 1 - Frequenza risposta al quesito 1.

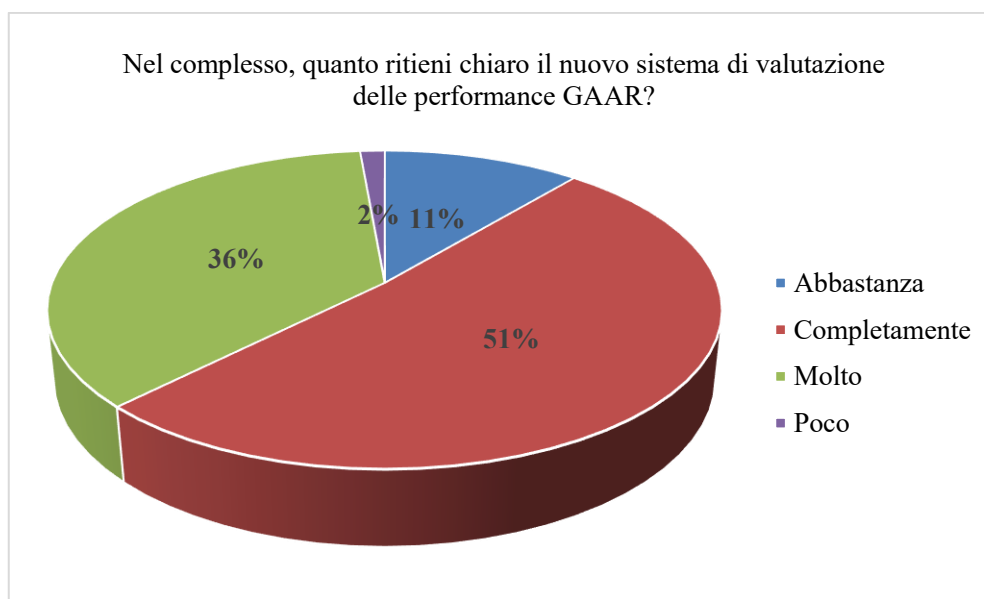


Figura 4 Grafico 2 - Frequenza risposta al quesito 2.

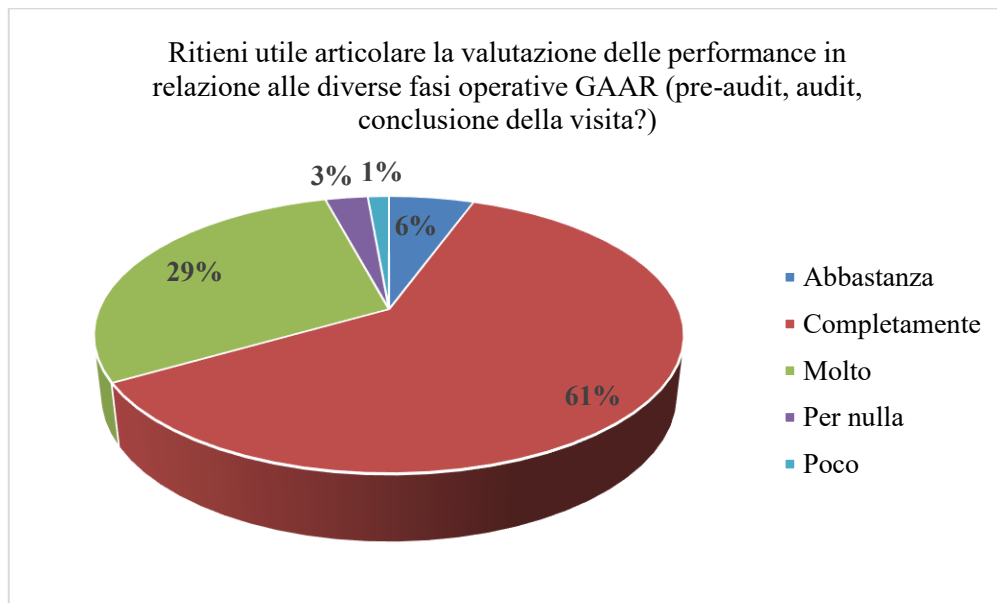


Figura 5 Grafico 3 - Frequenza risposta al quesito 3.

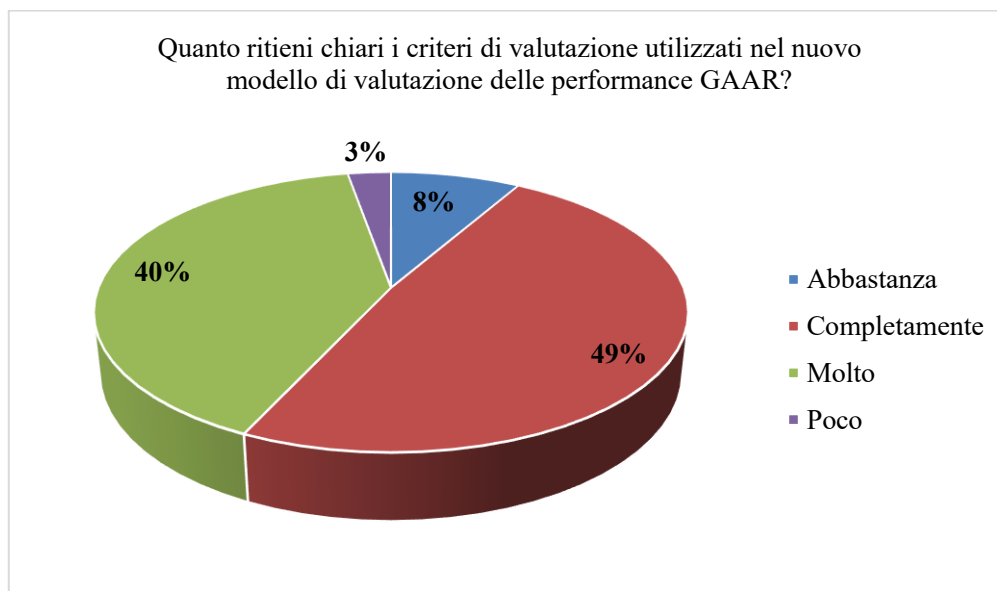


Figura 6 Grafico 4 - Frequenza risposta al quesito 4.

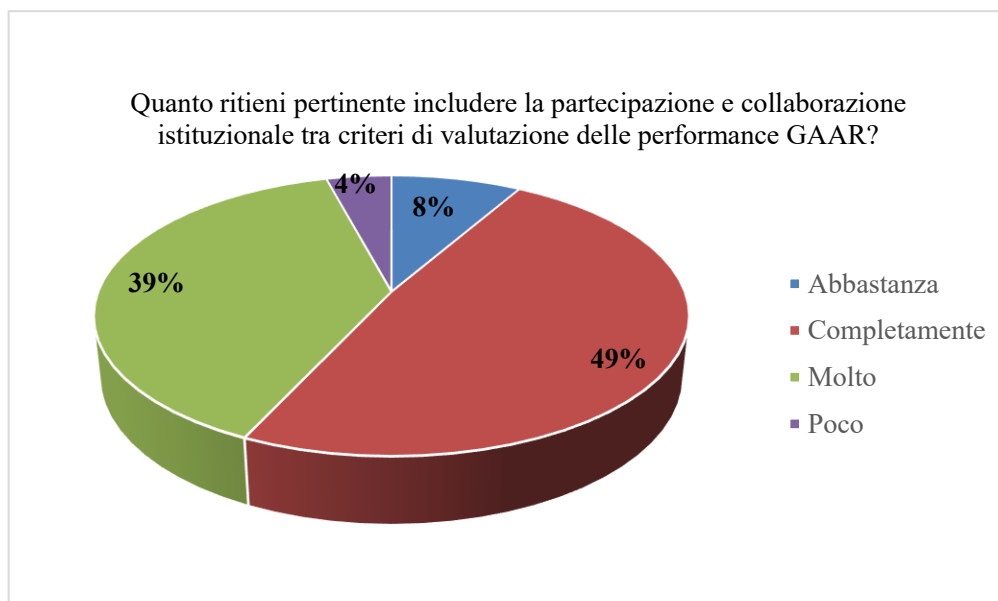


Figura 7 Grafico 5 - Frequenza risposta al quesito 5.

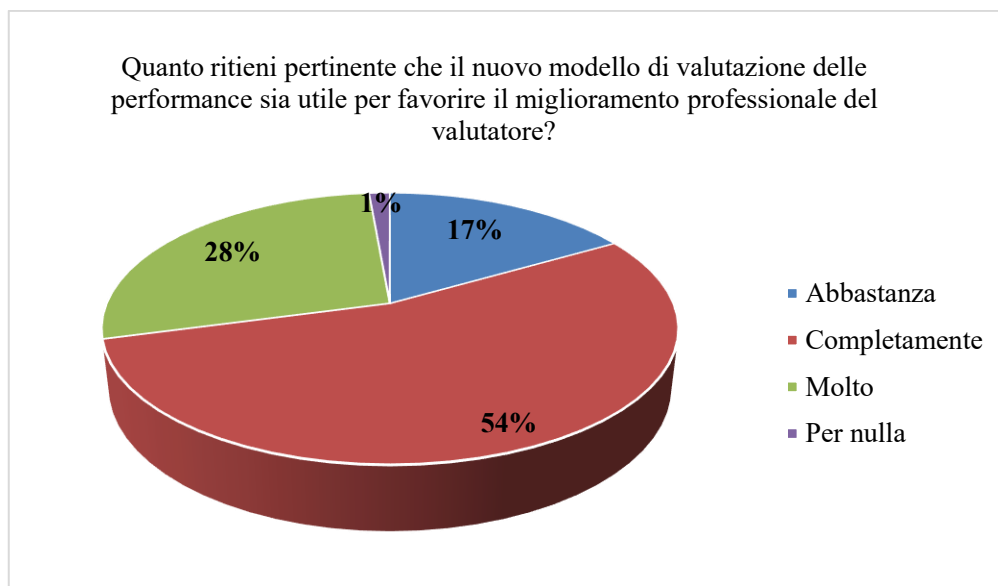


Figura 8 Grafico 6 - Frequenza risposta al quesito 6.

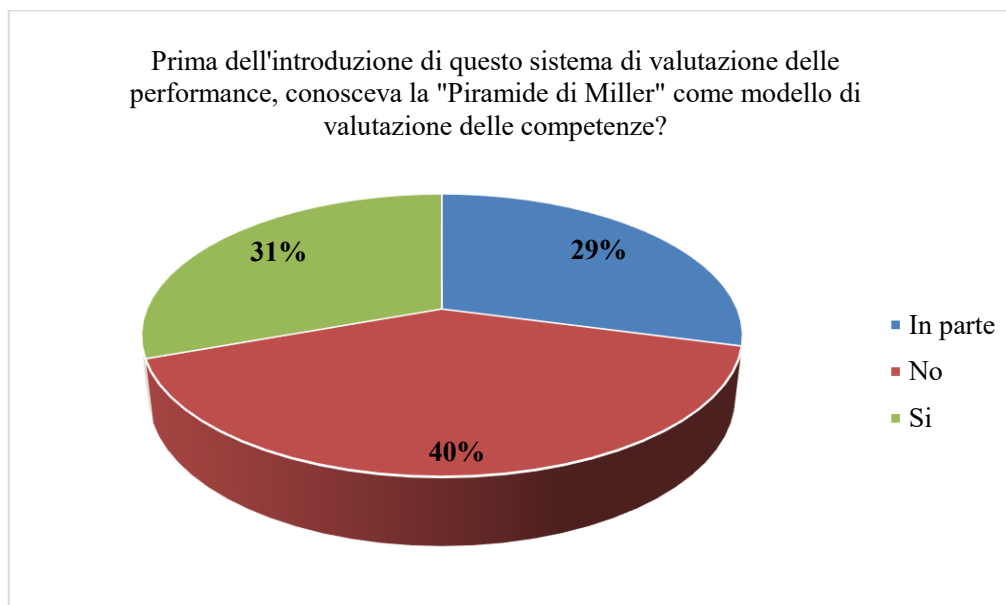
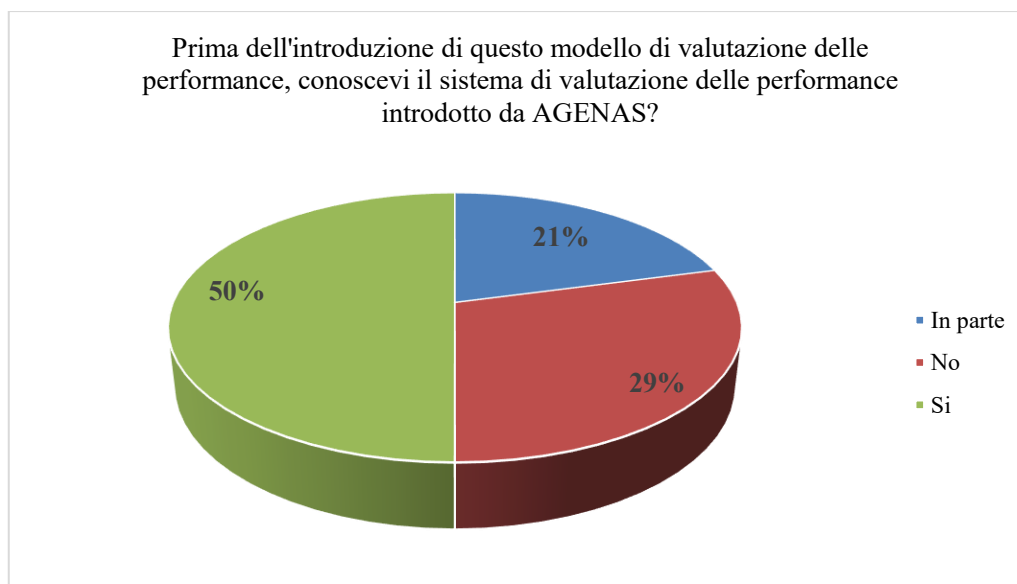


Figura 9 Grafico 7 - Frequenza risposta al quesito 7.



Le evidenze derivanti dall'analisi qualitativa delle risposte aperte sono state sintetizzate successivamente tramite analisi statistica e riportate nella Tabella 11. Considerata la natura non mutuamente esclusiva della codifica tematica, che ha consentito l'attribuzione

della medesima risposta a più categorie di analisi, si è ritenuto metodologicamente più appropriato presentare i risultati in forma tabellare piuttosto che mediante rappresentazione grafica, al fine di evitare possibili interpretazioni fuorvianti della distribuzione delle frequenze.

ANALISI STATISTICA

Il dataset analizzato comprende 73 rispondenti. Considerando come popolazione eleggibile i 115 valutatori GAAR indicati nella traccia metodologica, il tasso di risposta risulta pari al 63.5%. Lo score composito di percezione, ottenuto mediante il calcolo della media dei cinque item su scala Likert (Q1–Q5), ha evidenziato un livello complessivamente elevato di accettabilità del sistema valutativo introdotto (media = 4,36; deviazione standard = 0,68; IC95%: 4,20–4,52; mediana = 4,6). In particolare, il 76,7% dei partecipanti ha riportato un punteggio medio pari o superiore a 4/5, mentre il 50,7% ha raggiunto valori $\geq 4,5/5$, indicando una marcata prevalenza di valutazioni favorevoli. L'analisi dell'affidabilità interna dello strumento ha evidenziato un'eccellente consistenza psicometrica, con un coefficiente α di Cronbach pari a 0,900, suggerendo un'elevata omogeneità tra gli item e una buona capacità dello strumento di misurare in maniera coerente il costrutto sottostante. L'analisi fattoriale esplorativa ha ulteriormente supportato la validità strutturale del questionario, evidenziando una configurazione sostanzialmente monodimensionale dello strumento. In particolare, il primo fattore ha mostrato un autovalore pari a 3,59, spiegando il 71,9% della varianza totale, a sostegno dell'ipotesi che gli item concorrano prevalentemente alla misurazione di un unico costrutto latente riferibile alla percezione del nuovo sistema di valutazione delle performance.

Analisi descrittiva univariata degli item Likert

Tutti gli item hanno mostrato livelli medi elevati, compresi tra 4,32 e 4,44 su scala 1–5. L'item con punteggio medio più elevato è risultato Q2, relativo all'utilità dell'articolazione della valutazione nelle fasi operative, mentre il valore medio più contenuto, comunque elevato, è stato osservato per Q4, riferito all'adeguatezza della

dimensione partecipativa/collaborativa. La quota di risposte positive (Molto/Completamente) è stata superiore all'82% per tutti gli item.

Item	Dominio	Media	DS	Mediana	Q1	Q3	n positivo	% positivo	IC95% inf	IC95% sup
Q_1	Chiarezza complessiva del sistema di valutazione	4.356	0.752	5.000	4.000	5.000	63	86.301	4.181	4.532
Q_2	Utilità dell'articolazione per fasi operative	4.438	0.882	5.000	4.000	5.000	66	90.411	4.233	4.644
Q_3	Comprensibilità dei criteri di valutazione	4.342	0.749	4.000	4.000	5.000	65	89.041	4.168	4.517
Q_4	Adeguatezza della dimensione partecipativa/collaborativa	4.315	0.797	4.000	4.000	5.000	64	87.671	4.129	4.501
Q_5	Contributo al miglioramento professionale	4.329	0.851	5.000	4.000	5.000	60	82.192	4.130	4.527
<i>Nota. Risposte positive = Molto + Completamente; Q1 e Q3 indicano i quartili dello score dell'item.</i>										

Tabella 6 Statistiche descrittive degli item Q1-Q5

Item	Dominio	Per nulla	Poco	Abbastanza	Molto	Completamente
Q_1	Chiarezza complessiva del sistema di valutazione	0 (0.0%)	1 (1.4%)	9 (12.3%)	26 (35.6%)	37 (50.7%)
Q_2	Utilità dell'articolazione per fasi operative	2 (2.7%)	1 (1.4%)	4 (5.5%)	22 (30.1%)	44 (60.3%)
Q_3	Comprensibilità dei criteri di valutazione	0 (0.0%)	2 (2.7%)	6 (8.2%)	30 (41.1%)	35 (47.9%)
Q_4	Adeguatezza della dimensione partecipativa/collaborativa	0 (0.0%)	3 (4.1%)	6 (8.2%)	29 (39.7%)	35 (47.9%)
Q_5	Contributo al miglioramento professionale	1 (1.4%)	0 (0.0%)	12 (16.4%)	21 (28.8%)	39 (53.4%)

Tabella 7 Distribuzione percentuale delle risposte agli item Likert

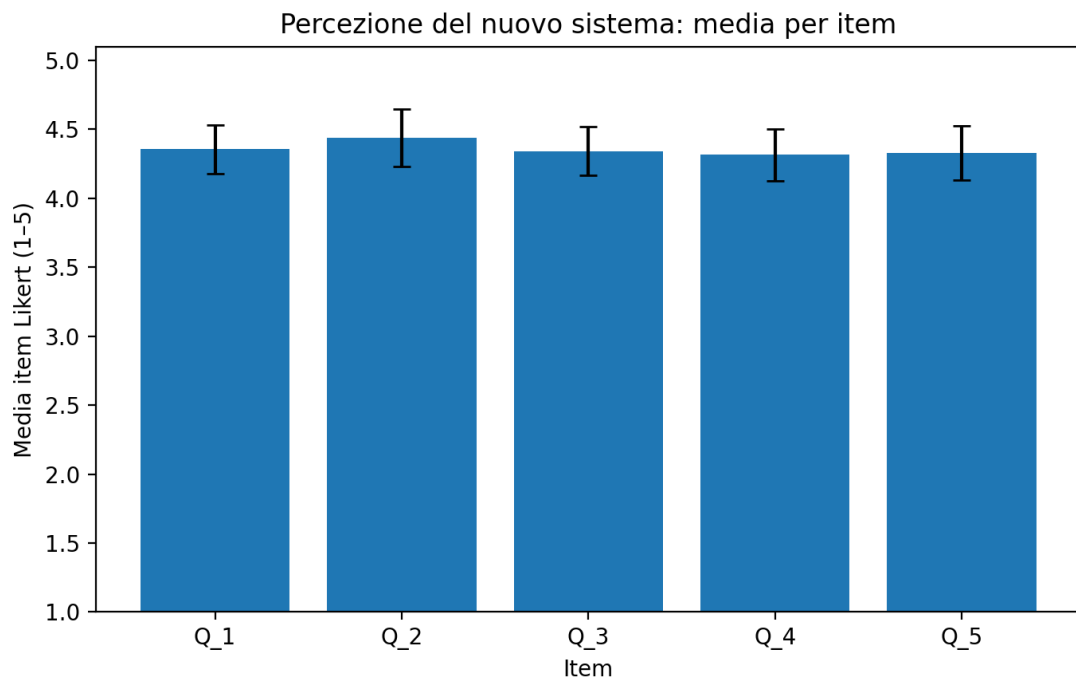


Figura 10 Media degli item Likert con intervallo di confidenza al 95%

Score composito di percezione/accettabilità

Lo score composito ha mostrato una media di 4.36 (DS=0.68; IC95% 4.20–4.52), con mediana pari a 4.6 e IQR 4.0–5.0. Il valore minimo osservato è stato 2.0 e il massimo 5.0. La distribuzione è risultata significativamente non normale (Shapiro-Wilk $W=0.862$; $p<0,001$), con evidente asimmetria verso valori elevati e possibile effetto ceiling. Il test di Wilcoxon rispetto al valore neutro 3/5 ha confermato una percezione significativamente superiore al livello intermedio della scala ($W=2604.5$; $p<0,001$).

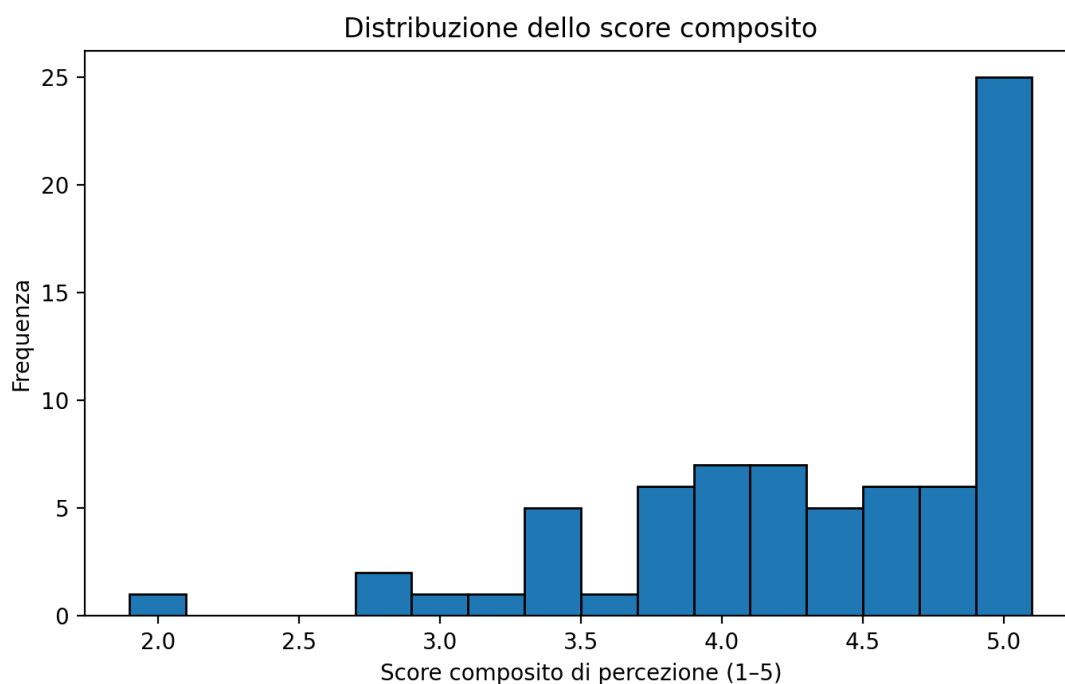


Figura 11 Distribuzione dello score composito di percezione

Item	Dominio	Mediana	Wilcoxon W	p
Q_1	Chiarezza complessiva del sistema di valutazione	5.000	2066.000	<0,001
Q_2	Utilità dell'articolazione per fasi operative	5.000	2310.000	<0,001
Q_3	Comprensibilità dei criteri di valutazione	4.000	2245.000	<0,001
Q_4	Adeguatezza della dimensione partecipativa/collaborativa	4.000	2228.500	<0,001
Q_5	Contributo al miglioramento professionale	5.000	1849.500	<0,001
Nota. Tutti i test sono monodirezionali verso valori superiori al punto neutro.				

Tabella 8 Test di Wilcoxon degli item rispetto al valore neutro della scala (3=abbastanza)

Affidabilità interna e dimensionalità dello strumento

La consistenza interna dei cinque item è risultata eccellente (α di Cronbach=0.900). Le correlazioni item-totale corrette sono risultate comprese tra 0.684 e 0.826; inoltre, l'eliminazione di ciascun item non ha prodotto incrementi sostanziali dell' α , confermando il contributo coerente di tutti gli item alla scala complessiva.

Item	Dominio	Correlazione item-totale corretta (rho)	Alpha se item eliminato
Q_1	Chiarezza complessiva del sistema di valutazione	0.729	0.882
Q_2	Utilità dell'articolazione per fasi operative	0.684	0.898
Q_3	Comprensibilità dei criteri di valutazione	0.825	0.872
Q_4	Adeguatezza della dimensione partecipativa/collaborativa	0.826	0.878
Q_5	Contributo al miglioramento professionale	0.818	0.857
<i>Nota. Correlazioni item-totale corrette calcolate con rho di Spearman.</i>			

Tabella 9 Consistenza interna dello strumento

L'adeguatezza campionaria per l'analisi fattoriale è risultata buona ($KMO=0.841$) e il test di sfericità di Bartlett è risultato significativo ($\chi^2=226.33$; $gdl=10.0$; $p<0,001$), indicando l'idoneità della matrice di correlazione. L'analisi degli autovalori ha evidenziato un primo fattore dominante (autovalore=3.59), in grado di spiegare il 71.9% della varianza totale; i restanti autovalori sono risultati inferiori a 1. I carichi fattoriali sul primo fattore sono risultati elevati per tutti gli item, supportando una struttura sostanzialmente monodimensionale dello strumento.

Item	Dominio	Carico fattoriale F1
Q_1	Chiarezza complessiva del sistema di valutazione	0.837
Q_2	Utilità dell'articolazione per fasi operative	0.778
Q_3	Comprensibilità dei criteri di valutazione	0.872
Q_4	Adeguatezza della dimensione partecipativa/collaborativa	0.841
Q_5	Contributo al miglioramento professionale	0.906
<i>Nota. Carichi fattoriali riportati in valore assoluto.</i>		

Tabella 10 Analisi fattoriale esplorativa: carichi sul primo fattore

Item	Q_1	Q_2	Q_3	Q_4	Q_5
Q_1	1.000	0.608	0.734	0.617	0.694
Q_2	0.608	1.000	0.622	0.629	0.602
Q_3	0.734	0.622	1.000	0.757	0.748
Q_4	0.617	0.629	0.757	1.000	0.765
Q_5	0.694	0.602	0.748	0.765	1.000
<i>Nota. Tutte le correlazioni tra item sono risultate statisticamente significative ($p < 0,001$).</i>					

Tabella 11 Matrice di correlazione Spearman tra item Q1-Q5

Familiarità con i modelli di riferimento e analisi bivariata

La familiarità con la Piramide di Miller è risultata completa nel 30.1% dei rispondenti, parziale nel 28.8% e assente nel 41.1%. La familiarità con il sistema AGENAS è risultata completa nel 49.3%, parziale nel 20.5% e assente nel 30.1%. La familiarità con i due modelli è risultata significativamente associata ($\chi^2=23.52$; $gdl=4$; $p<0,001$), suggerendo una convergenza tra conoscenza del framework teorico e conoscenza del sistema istituzionale di riferimento.

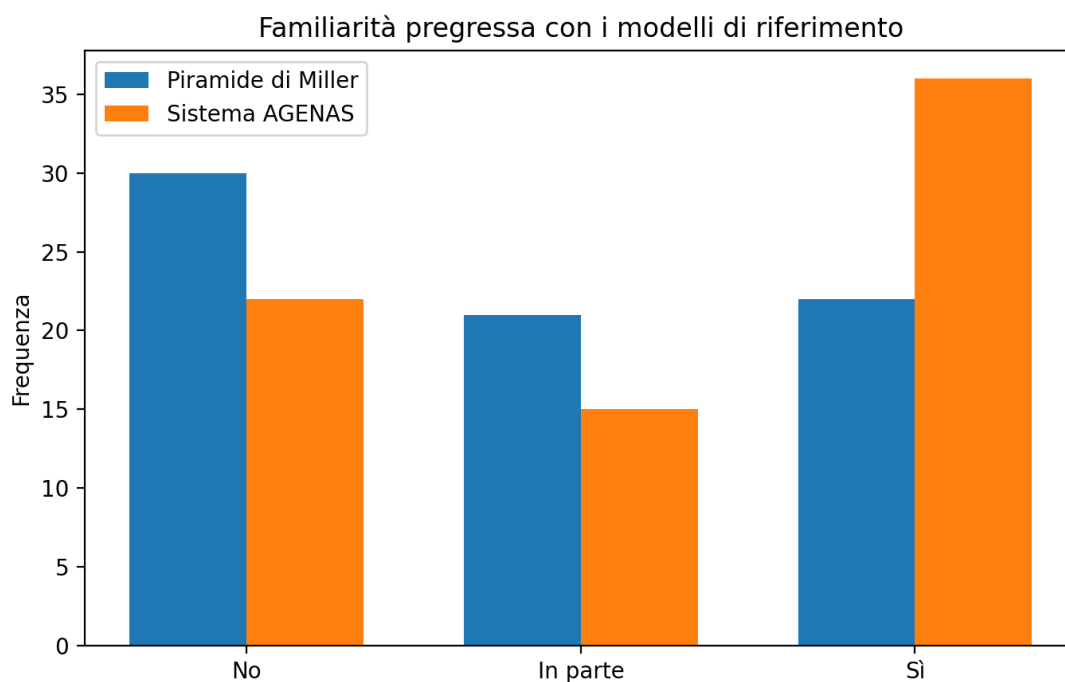


Figura 12 Distribuzione della familiarità con Piramide di Miller e sistema AGENAS

Familiarità Miller	n.	Media score	DS	Mediana
No	30	4.200	0.701	4.200
In parte	21	4.476	0.504	4.600
Si	22	4.455	0.786	4.900
<i>Nota. Kruskal-Wallis $H=3.77$; $p=0.152$.</i>				

Tabella 12 Score composito per familiarità con la Piramide di Miller

Familiarità AGENAS	n.	Media score	DS	Mediana
No	22	4.209	0.716	4.300
In parte	15	4.333	0.653	4.400
Si	36	4.456	0.675	4.600
<i>Nota. Kruskal-Wallis $H=2.32$; $p=0.313$.</i>				

Tabella 13 Score composito per familiarità con il sistema AGENAS

Le differenze dello score composito tra i livelli di familiarità non sono risultate statisticamente significative né per la Piramide di Miller ($H=3.77$; $p=0.152$) né per il sistema AGENAS ($H=2.32$; $p=0.313$). Anche le correlazioni ordinali tra familiarità e score sono risultate deboli: $\rho=0.227$ per Miller ($p=0.054$) e $\rho=0.180$ per AGENAS ($p=0.128$).

Miller / AGENAS	No	In parte	Si
No	17	7	6
In parte	2	6	13
Si	3	2	17
<i>Nota. Test χ^2 su tabella 3×3.</i>			

Tabella 14 Associazione tra familiarità con Piramide di Miller e sistema AGENAS

Analisi multivariata: regressione lineare esplorativa

Nel modello di regressione lineare multipla, la familiarità con la Piramide di Miller e con il sistema AGENAS non sono risultate predittori statisticamente significativi dello score composito. Il modello ha spiegato una quota limitata della variabilità dello score ($R^2=0.035$; R^2 corretto=0.007). L'assenza di significatività, in presenza di punteggi complessivamente elevati e compressi verso il valore massimo, suggerisce che la percezione positiva del nuovo sistema sia ampiamente diffusa tra i rispondenti, indipendentemente dalla familiarità pregressa con i due modelli teorico-istituzionali.

Variabile	B	SE robusta HC3	IC95% inf	IC95% sup	p
Intercetta	4.180	0.152	3.877	4.482	<0,001
Familiarità Piramide di Miller (0=No, 1=In parte, 2=Si)	0.093	0.101	-0.108	0.295	0.359
Familiarità sistema AGENAS (0=No, 1=In parte, 2=Si)	0.078	0.087	-0.095	0.252	0.371
<i>Nota. Errori standard robusti HC3; $R^2=0.035$; VIF predittori=1.33 e 1.33.</i>					

Tabella 15 Regressione lineare multipla sullo score composito di percezione

Analisi qualitativa della domanda aperta

Le risposte aperte sono state sottoposte a una content analysis descrittiva, con codifica tematica non mutuamente esclusiva. La maggioranza dei rispondenti non ha formulato proposte di modifica o ha confermato l'adeguatezza del modello. Le aree di miglioramento più ricorrenti hanno riguardato l'affinamento di criteri/metodologie/indicatori, la valorizzazione di momenti di confronto e feedback, e l'introduzione di logiche di monitoraggio longitudinale, formazione mirata e miglioramento continuo.

Tema emerso	n	% su rispondenti
Assenza di suggerimenti o conferma del modello	57	78.082
Valutazione positiva/adeguatezza percepita	16	21.918
Affinamento di criteri/metodologia/indicatori	9	12.329
Confronto, feedback e partecipazione del valutato	7	9.589
Monitoraggio longitudinale/formazione/miglioramento continuo	5	6.849
Benchmarking esterno	1	1.370
<i>Nota. Categorie non mutuamente esclusive; una risposta poteva contribuire a più temi.</i>		

Tabella 16 Analisi tematica delle risposte aperte

DISCUSSIONE

La discussione dei risultati è stata sviluppata attraverso un'analisi critica dei principali riferimenti teorici e istituzionali che hanno costituito il fondamento metodologico del presente studio. In particolare, vengono approfonditi il Sistema di Misurazione e Valutazione della Performance di AGENAS e la Piramide di Miller, modelli che hanno orientato la progettazione della nuova scheda di valutazione. Successivamente, i risultati della sperimentazione vengono interpretati alla luce delle evidenze emerse, al fine di evidenziare il contributo del nuovo strumento al processo di valutazione delle performance dei valutatori GAAR e le possibili implicazioni per il sistema di autorizzazione e accreditamento sanitario.

SISTEMA DI MISURAZIONE E VALUTAZIONE DELLE PERFORMANCE AGENAS – Agenzia nazionale per i servizi sanitari regionali

L'Agenzia Nazionale per i Servizi Sanitari Regionali (AGENAS) è un ente pubblico non economico di rilievo nazionale, istituito con decreto legislativo 30 giugno 1993, n. 266 e successive modificazioni. Essa opera quale organo tecnico-scientifico del Servizio Sanitario Nazionale e, ai sensi dell'art. 2, comma 357, della legge 24 dicembre 2007, n. 244, svolge attività di ricerca, consulenza, formazione e supporto nei confronti del Ministero della Salute, delle Regioni e delle Province Autonome di Trento e Bolzano. Dotata di autonomia statutaria, regolamentare, organizzativa, amministrativa, finanziaria e contabile, l'Agenzia è sottoposta alla vigilanza del Ministero della Salute e al controllo della Corte dei Conti [12]. I compiti istituzionali sono individuati dalla Conferenza Unificata ai sensi dell'art. 9, comma 2, lett. g), del decreto legislativo 28 agosto 1997, n. 281, oltre che dalle ulteriori disposizioni normative vigenti. Tra le finalità prioritarie dell'AGENAS rientra il supporto tecnico-operativo alle politiche statali e regionali di governo del sistema sanitario, con particolare riferimento all'organizzazione dei servizi, alla qualità delle prestazioni assistenziali e ai processi formativi. Tale obiettivo viene perseguito anche mediante attività di ricerca, monitoraggio e valutazione orientate allo sviluppo e al miglioramento del sistema salute. In questo contesto, l'Agenzia ha progressivamente implementato il processo di misurazione e valutazione della performance, attraverso l'aggiornamento annuale del Sistema di misurazione e

valutazione della performance, al fine di assicurare coerenza con il quadro normativo vigente e rispondere in maniera più efficace alle esigenze organizzative interne. Il Sistema disciplina fasi, soggetti coinvolti, responsabilità, modalità e tempistiche del processo valutativo, integrandosi con gli strumenti di programmazione finanziaria e di bilancio. Esso consente la misurazione di parametri quantitativi e semi-quantitativi, finalizzati alla valutazione della performance strategica dell'Agenzia, della performance organizzativa delle singole strutture e del contributo individuale del personale. La performance viene intesa quale contributo apportato dall'organizzazione nel suo complesso, nonché dalle singole unità operative e dai dipendenti, al raggiungimento degli obiettivi istituzionali programmati. In tale prospettiva, l'Agenzia programma, monitora, misura e rende pubblici i risultati conseguiti ai diversi livelli organizzativi, promuovendo trasparenza, responsabilità e apprendimento organizzativo, in coerenza con il principio di accountability. Il quadro normativo di riferimento è rappresentato, in particolare, dall'art. 7 del decreto legislativo n. 150/2009, come modificato dal d.lgs. n. 74/2017, che impone alle amministrazioni pubbliche l'adozione e l'aggiornamento annuale del Sistema di misurazione e valutazione della performance, previo parere vincolante dell'Organismo Indipendente di Valutazione (OIV). Inoltre, l'art. 45 del d.lgs. n. 165/2001 collega la performance sia all'amministrazione nel suo complesso sia alle singole unità organizzative o aree di responsabilità. Dal punto di vista metodologico, il Sistema è sviluppato nel rispetto del quadro normativo delineato dal d.lgs. n. 150/2009 e successive modificazioni, nonché dei criteri interpretativi contenuti nelle delibere della ex CIVIT e nelle Linee Guida del Dipartimento della Funzione Pubblica, considerate di carattere generale e applicabili, in assenza di specifiche indicazioni, anche alle altre amministrazioni dello Stato. Il Sistema di Misurazione e Valutazione della Performance adottato da AGENAS si fonda su una concezione integrata della performance, articolata su due dimensioni strettamente interconnesse: quella organizzativa e quella individuale. Tale impostazione consente di ricondurre l'azione delle singole strutture e dei singoli dipendenti all'interno di un quadro coerente con la mission istituzionale dell'Ente e con gli indirizzi strategici definiti a livello centrale. La performance organizzativa è intesa come la capacità dell'Agenzia e delle sue articolazioni interne di perseguire gli obiettivi strategici programmati, assicurando coerenza tra attività svolte, risultati conseguiti e risorse impiegate. Accanto alla dimensione organizzativa, la performance individuale

misura il contributo del singolo al conseguimento degli obiettivi istituzionali e si basa su una duplice componente: il raggiungimento di obiettivi individuali e la valutazione dei comportamenti organizzativi. Gli obiettivi individuali sono definiti in coerenza con la programmazione annuale e possono riguardare, ad esempio, il rispetto delle tempistiche del ciclo della performance, l'adempimento degli obblighi in materia di trasparenza e protezione dei dati personali, l'attuazione delle misure di prevenzione della corruzione e il rispetto di specifici target operativi. Tali obiettivi sono associati a indicatori misurabili e a percentuali di raggiungimento che determinano l'attribuzione del punteggio. Particolare rilievo assume la valutazione dei comportamenti organizzativi, che integra la dimensione quantitativa dei risultati con quella qualitativa delle competenze professionali. Gli item considerati includono, tra gli altri, la capacità di programmazione delle attività nel rispetto degli indirizzi strategici, il monitoraggio e il controllo delle azioni intraprese, la leadership e la capacità di coordinamento, la promozione del benessere organizzativo, la collaborazione e l'integrazione con il gruppo di lavoro, nonché la capacità di differenziare le valutazioni e di garantire il rispetto degli adempimenti normativi. La valutazione di tali comportamenti avviene attraverso schede strutturate che attribuiscono punteggi su scale predefinite e confluisce in una valutazione finale annuale. La misurazione del grado di realizzazione delle attività è associata a giudizi qualitativi (quali "eccezionale", "pienamente soddisfacente", "soddisfacente", "migliorabile", "non soddisfacente") e incide sulla determinazione della premialità, secondo una logica di valorizzazione del merito e di responsabilizzazione diffusa. L'intero sistema è strutturato in modo da garantire coerenza, prevedendo momenti di confronto tra valutatore e valutato e procedure di riesame a più livelli, al fine di assicurare imparzialità, trasparenza e correttezza del processo valutativo. **In tal modo, la misurazione della performance non si limita a una funzione di controllo, ma si configura come strumento di apprendimento organizzativo, di orientamento dei comportamenti e di rafforzamento dell'accountability istituzionale [9].**

LA PIRAMIDE DI MILLER

La **Piramide di Miller**, (Figura 14) elaborata nel 1990 da **George E. Miller**, medico ed esperto statunitense di pedagogia medica, rappresenta uno dei modelli teorici più autorevoli per la valutazione delle competenze professionali in ambito sanitario. Miller,

docente presso l'University of Illinois e figura centrale nello sviluppo della medical education moderna, ha contribuito in modo determinante al superamento di modelli valutativi esclusivamente nozionistici, promuovendo un approccio orientato alla misurazione delle competenze nella pratica reale. Il modello distingue quattro livelli progressivi di competenza: *"knows"* (conoscenza teorica), *"knows how"* (capacità di applicare le conoscenze), *"shows how"* (dimostrazione operativa della competenza) e *"does"* (agire competente nella pratica quotidiana). Tale articolazione consente di strutturare la valutazione non soltanto sulla base del sapere dichiarato, ma soprattutto sull'osservazione delle performance effettivamente agite nel contesto professionale [10]. In ambito sanitario, l'applicazione della Piramide di Miller permette di integrare indicatori quantitativi di risultato con dimensioni qualitative relative ai comportamenti organizzativi, alla capacità decisionale e alla responsabilità professionale. In questa prospettiva, il modello si configura come uno strumento metodologico particolarmente idoneo alla costruzione di sistemi di valutazione del personale che valorizzino sia il raggiungimento degli obiettivi sia la qualità dell'azione professionale concretamente esercitata, contribuendo al rafforzamento dell'accountability e del miglioramento continuo [13]. Ulteriore elemento di rilevanza del modello risiede nella sua ampia diffusione e validazione nell'ambito della formazione e della valutazione delle professioni sanitarie a livello internazionale. La Piramide di Miller è infatti riconosciuta come uno dei principali riferimenti teorici nei sistemi di assessment competency-based [14], in quanto consente di superare approcci valutativi limitati alla sola verifica delle conoscenze teoriche, orientando l'attenzione verso la capacità del professionista di trasferire e applicare tali conoscenze nella pratica clinica reale. Dal punto di vista metodologico, il modello introduce una progressione gerarchica che permette di distinguere tra diversi livelli di complessità della competenza, favorendo l'utilizzo di strumenti valutativi differenziati e coerenti con il livello da indagare. In particolare, mentre i livelli inferiori (*"knows"* e *"knows how"*) possono essere valutati attraverso strumenti tradizionali quali test scritti o prove teoriche, i livelli superiori (*"shows how"* e *"does"*) richiedono metodologie più complesse, basate sull'osservazione diretta, sulla simulazione e sull'analisi delle performance in contesti reali. Tale impostazione rende la Piramide di Miller particolarmente rilevante nei contesti in cui la valutazione deve tener conto non solo delle conoscenze possedute, ma anche della capacità di agire in modo

efficace, appropriato e responsabile all'interno di sistemi organizzativi complessi [15]. In questo senso, il modello si configura come un riferimento teorico solido per la progettazione di sistemi di valutazione orientati alla misurazione della competenza nella sua dimensione più completa, integrando sapere, saper fare e comportamento professionale [16].

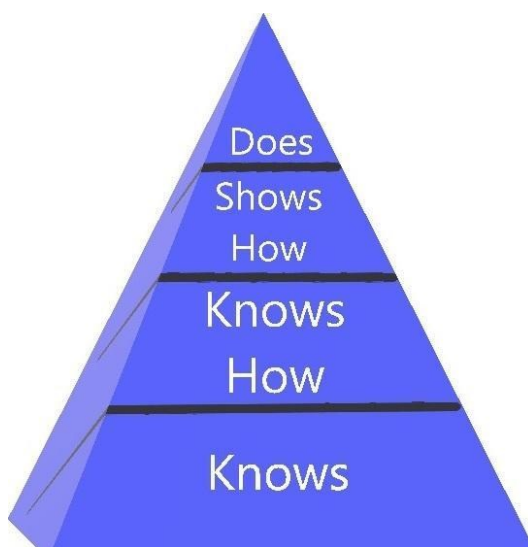


Figura 13 La Piramide di Miller

VALUTAZIONE CRITICA DEL SISTEMA PRECEDENTEMENTE IN USO

L'analisi critica del sistema di valutazione precedentemente adottato presso l'Organismo Tecnicamente Accreditante (OTA) della Regione Marche ha evidenziato la presenza di uno strumento formalizzato e già orientato alla misurazione di diversi aspetti dell'attività svolta dai valutatori GAAR. La scheda risultava infatti articolata in una serie di item riferiti a specifiche dimensioni operative, quali il rispetto delle tempistiche, la partecipazione alle attività istituzionali, la qualità della documentazione prodotta e la corretta esecuzione delle attività di verifica, utilizzando una scala di valutazione di tipo Likert a dieci livelli. Pur rappresentando un importante tentativo di standardizzazione del processo valutativo, il sistema presentava alcune criticità metodologiche che ne limitavano la capacità di fornire una valutazione pienamente strutturata delle competenze professionali del valutatore. In primo luogo, non risultava esplicitato alcun framework

teorico di riferimento a supporto della costruzione della scheda, rendendo meno evidente il razionale metodologico sottostante alla selezione degli item e alla definizione delle dimensioni valutative considerate.

Un ulteriore elemento di criticità era rappresentato dalla prevalente focalizzazione sugli aspetti procedurali e organizzativi dell'attività del GAAR. Sebbene tali dimensioni costituiscano componenti essenziali della performance professionale, esse non consentono, da sole, di cogliere in maniera esaustiva la complessità delle competenze richieste al valutatore, quali la capacità di interpretazione dei requisiti, il ragionamento professionale, il giudizio tecnico, l'applicazione delle conoscenze normative e la gestione delle dinamiche relazionali durante l'attività di audit. Anche la struttura della scala di valutazione presentava alcuni potenziali limiti. L'impiego di una scala Likert a dieci punti, in assenza di descrittori operativi chiaramente definiti per ciascun livello di risposta, poteva infatti favorire un'elevata discrezionalità nell'attribuzione dei punteggi, con il rischio di incrementare la variabilità inter-valutatore e ridurre l'uniformità interpretativa del processo di valutazione. Ulteriori criticità riguardavano la limitata articolazione della valutazione nelle diverse fasi del percorso ispettivo e la mancanza di una distinzione strutturata tra competenze tecnico-professionali, comportamenti organizzativi e partecipazione istituzionale. Tale impostazione poteva rendere meno agevole identificare con precisione le aree di forza e le eventuali criticità del singolo valutatore, limitando conseguentemente anche il potenziale utilizzo della valutazione quale strumento di sviluppo professionale e miglioramento continuo. Alla luce di tali considerazioni, la nuova scheda di valutazione sviluppata nel presente studio è stata progettata con l'obiettivo di superare tali limiti, attraverso l'integrazione di un modello teorico validato quale la Piramide di Miller con il Sistema di Misurazione e Valutazione della Performance di AGENAS. L'introduzione di criteri maggiormente osservabili, indicatori esplicitamente definiti, una diversa articolazione delle dimensioni valutative e una più chiara strutturazione del processo nelle fasi di pre-audit, audit e conclusione della visita ha consentito di sviluppare uno strumento orientato non soltanto alla misurazione degli adempimenti procedurali, ma anche alla valutazione delle competenze professionali e del contributo organizzativo del valutatore. In tale prospettiva, il passaggio dal sistema precedentemente in uso al modello "Miller-AGENAS" rappresenta un significativo avanzamento metodologico verso una valutazione maggiormente oggettiva,

standardizzata, riproducibile e coerente con i principi della moderna valutazione delle performance in ambito sanitario.

SIGNIFICATO METODOLOGICO DELL'INTEGRAZIONE TRA PIRAMIDE DI MILLER E SISTEMA AGENAS

Uno degli elementi di maggiore rilevanza emersi dal presente studio è rappresentato dalla possibilità di integrare un modello teorico di valutazione delle competenze, quale la Piramide di Miller, con un sistema istituzionale di misurazione della performance, come quello sviluppato da AGENAS. La Piramide di Miller consente infatti di strutturare la valutazione su livelli progressivi di competenza, superando un approccio esclusivamente nozionistico e orientandosi verso l'osservazione delle performance nella pratica reale ("does"). Tale impostazione risulta particolarmente coerente con le caratteristiche del ruolo dei valutatori GAAR, la cui attività si esplica prevalentemente in contesti operativi complessi, quali le visite ispettive, la valutazione documentale e la redazione dei verbali di audit. D'altro canto, il sistema AGENAS introduce una dimensione organizzativa e sistemica della valutazione, fondata su indicatori di performance, accountability e trasparenza, in linea con quanto previsto dal D.Lgs. 150/2009. L'integrazione di tali due modelli ha consentito di sviluppare uno strumento valutativo in grado di coniugare la dimensione individuale della competenza con quella organizzativa della performance, garantendo al contempo coerenza metodologica e applicabilità operativa. Questo approccio integrato rappresenta un elemento innovativo rispetto ai sistemi di valutazione tradizionali, che spesso tendono a focalizzarsi esclusivamente su indicatori quantitativi o su aspetti procedurali, trascurando la complessità delle competenze professionali agite.

INTERPRETAZIONE DEI RISULTATI DEL QUESTIONARIO DI GRADIMENTO

L'analisi preliminare qualitativa delle risposte alla domanda aperta ha consentito di individuare diverse aree tematiche, che forniscono ulteriori elementi interpretativi rispetto ai dati quantitativi. In primo luogo, emerge una diffusa percezione di adeguatezza dello strumento, ritenuto coerente con le attività svolte e in grado di rappresentare in maniera significativa le competenze professionali dei valutatori. In secondo luogo, sono stati rilevati suggerimenti di miglioramento tecnico, in particolare relativi alla necessità di

rendere ancora più chiari alcuni criteri valutativi e di definire in maniera più puntuale alcuni indicatori. Questo elemento evidenzia una partecipazione attiva dei valutatori al processo di miglioramento dello strumento e rappresenta un indicatore positivo di coinvolgimento organizzativo. Un ulteriore aspetto emerso riguarda la possibile modulazione del peso delle diverse dimensioni valutate in funzione del ruolo e delle attività svolte, suggerendo l'opportunità di sviluppare, in futuro, sistemi di valutazione più flessibili e personalizzati. Infine, una quota residuale di risposte si caratterizza per contenuti sintetici o non rilevanti ai fini dell'analisi, elemento fisiologico nei questionari a risposta aperta. Al fine di garantire una lettura sistematica e metodologicamente rigorosa dei risultati emersi, i dati raccolti attraverso il questionario di gradimento sono stati elaborati e rappresentati mediante specifiche restituzioni grafiche, funzionali a evidenziare la distribuzione delle risposte per ciascun item indagato. L'analisi dei grafici consente di approfondire la percezione dei valutatori in relazione ai principali aspetti del nuovo sistema di valutazione, con particolare riferimento alla chiarezza, alla coerenza metodologica, all'utilità operativa e al contributo al miglioramento professionale. Nei paragrafi seguenti vengono pertanto presentati e interpretati i risultati preliminari, al fine di fornire una lettura integrata e critica dei dati raccolti.

Grafico 1: Chiarezza del sistema di valutazione

Quesito: “Nel complesso quanto ritieni chiaro il nuovo sistema di valutazione delle performance GAAR?”

Il grafico evidenzia la distribuzione delle risposte relative alla percezione di chiarezza del nuovo sistema di valutazione delle performance GAAR. I risultati mostrano una prevalenza significativa di giudizi positivi, con una concentrazione delle risposte nei livelli più elevati della scala Likert, “molto” 36% e “completamente” 51%. In particolare, il numero di rispondenti che ha espresso una valutazione “completamente” positiva risulta superiore rispetto alle altre categorie, indicando un elevato livello di comprensione del modello valutativo introdotto. Le risposte di livello intermedio “abbastanza” 11% risultano comunque presenti, suggerendo un buon grado generale di chiarezza, pur lasciando margini di miglioramento in termini di esplicitazione di alcuni elementi dello strumento. Le valutazioni negative “poco” 2% risultano residuali, a conferma dell'efficacia comunicativa del sistema.

Grafico 2: Utilità della valutazione nelle diverse fasi (pre-audit, audit, conclusione).

Quesito: “Ritieni utile articolare la valutazione delle performance in relazione alle diverse fasi operative GAAR (pre-audit. Audit, conclusione della visita)?”

Il grafico rappresenta la percezione di utilità dell’articolazione del sistema valutativo nelle diverse fasi operative del processo GAAR (pre-audit, audit e fase conclusiva). I risultati evidenziano una distribuzione fortemente orientata verso i livelli più elevati della scala Likert, “completamente” 61%, indicando come la strutturazione della valutazione nelle diverse fasi venga percepita come funzionale e coerente con le attività svolte. Tale dato suggerisce che la suddivisione del processo valutativo contribuisce a una maggiore aderenza tra criteri di valutazione e pratica operativa, migliorando la capacità dello strumento di rappresentare in modo accurato le competenze professionali.

Grafico 3: Chiarezza dei criteri di valutazione

Quesito: “Quanto ritieni chiari i criteri di valutazione utilizzati nel nuovo modello di valutazione delle performance GAAR?”

Il grafico relativo alla chiarezza dei criteri di valutazione evidenzia una prevalenza di risposte positive, con una concentrazione significativa nei livelli “molto” 40% e “completamente” 49%. Questo risultato indica che i criteri utilizzati risultano generalmente comprensibili e ben definiti. Tuttavia, la presenza di una quota di risposte collocate nei livelli intermedi suggerisce la possibilità di ulteriori miglioramenti nella formulazione degli indicatori, al fine di ridurre eventuali ambiguità interpretative e aumentare l’uniformità nell’applicazione del giudizio valutativo.

Grafico 4: Pertinenza della dimensione partecipativa/collaborativa

Quesito: “Quanto ritieni pertinente includere la partecipazione e collaborazione istituzionale tra criteri di valutazione delle performance GAAR?”

Il grafico evidenzia la percezione di pertinenza dell’inclusione della dimensione di partecipazione e collaborazione istituzionale all’interno del sistema di valutazione. I risultati mostrano una distribuzione fortemente orientata verso i livelli più elevati della scala Likert, “completamente” 49% e “molto” 39% indicando che tale dimensione è considerata rilevante e coerente con il ruolo dei valutatori GAAR. Questo dato sottolinea

l'importanza delle competenze relazionali e organizzative nel contesto dei processi di autorizzazione e accreditamento, confermando la validità dell'approccio multidimensionale adottato nella costruzione dello strumento valutativo.

Grafico 5: Contributo al miglioramento professionale

Quesito: “Quanto ritieni pertinente che il nuovo modello di valutazione delle performance sia utile per favorire il miglioramento professionale del valutatore?”

Il grafico relativo al contributo dello strumento al miglioramento professionale evidenzia una percezione complessivamente positiva da parte dei rispondenti. La maggior parte delle risposte si colloca nei livelli “molto” 28 % e “completamente” 54% , suggerendo che il nuovo sistema valutativo viene percepito non solo come strumento di misurazione, ma anche come leva di sviluppo delle competenze. Tale risultato è coerente con i principi di miglioramento continuo e formazione permanente che caratterizzano i sistemi di valutazione moderni, in particolare in ambito sanitario.

Grafico 6: Familiarità con la Piramide di Miller

Quesito: “Prima dell'introduzione di questo sistema di valutazione delle performance, conosceva la “Piramide di Miller” come modello di valutazione delle competenze?”

Il grafico evidenzia il livello di familiarità dei valutatori con il modello teorico della Piramide di Miller. I risultati mostrano una distribuzione dicotomica (sì/no), dalla quale emerge che una parte dei partecipanti non possedeva una conoscenza pregressa del modello “no” 40%, “in parte” 29%. Questo dato assume particolare rilevanza in quanto evidenzia come l'introduzione dello strumento valutativo abbia rappresentato anche un'opportunità di diffusione di modelli teorici avanzati, contribuendo ad arricchire il patrimonio culturale e metodologico dei valutatori.

Grafico 7: Familiarità con il sistema AGENAS

Quesito: “Prima dell'introduzione di questo modello di valutazione delle performance, conoscevi il sistema di valutazione delle performance introdotto da AGENAS?”

Il grafico relativo alla familiarità con il sistema di valutazione AGENAS mostra una distribuzione delle risposte che riflette una conoscenza non uniforme del modello tra i

partecipanti, “si” 50%, “no” 29%, “in parte” 21%. Questo elemento suggerisce la necessità di rafforzare le attività formative e informative sui sistemi istituzionali di valutazione della performance. L’integrazione del modello AGENAS nello strumento sviluppato rappresenta, in questo senso, un’opportunità per promuovere una maggiore diffusione dei principi di misurazione della performance nel contesto operativo dei valutatori GAAR.

INTERPRETAZIONE COMPLESSIVA DEI RISULTATI

La sperimentazione ha coinvolto una popolazione eleggibile di 115 valutatori GAAR, con 73 rispondenti al questionario di gradimento, pari a un tasso di risposta del 63,5%. Questo dato rappresenta un primo elemento rilevante, poiché documenta una partecipazione consistente da parte dei soggetti direttamente coinvolti nel processo valutativo e fornisce una base dati adeguata per una prima analisi descrittiva e psicometrica dello strumento. Sul piano della percezione complessiva, lo score composito, calcolato come media dei primi cinque item Likert, ha mostrato un valore medio pari a 4,36 su 5, con deviazione standard di 0,68, intervallo di confidenza al 95% compreso tra 4,20 e 4,52 e mediana pari a 4,6. Tali valori indicano una valutazione complessivamente molto favorevole del nuovo sistema. In particolare, il 76,7% dei partecipanti ha riportato uno score medio pari o superiore a 4/5, mentre il 50,7% ha raggiunto valori pari o superiori a 4,5/5. Questo andamento suggerisce una marcata accettabilità percepita dello strumento da parte dei valutatori coinvolti. L’analisi dei singoli item conferma tale orientamento positivo. Tutti i domini indagati hanno mostrato punteggi medi elevati, compresi tra 4,32 e 4,44. L’item con media più alta è risultato quello relativo all’utilità dell’articolazione della valutazione nelle diverse fasi operative, con media 4,438 e quota di risposte positive pari al 90,411%. Questo dato assume particolare rilievo perché conferma la coerenza della scelta progettuale di suddividere la valutazione secondo le fasi reali dell’attività GAAR, ovvero pre-audit, audit e conclusione della visita. Anche la comprensibilità dei criteri di valutazione ha ottenuto un risultato elevato, con media 4,342 e risposte positive pari all’89,041%, indicando che la strutturazione dello strumento in item, standard e criteri osservabili è stata percepita come chiara e utilizzabile. La chiarezza complessiva del sistema ha registrato una media pari a 4,356, con il 50,7% dei rispondenti che ha espresso il livello massimo “completamente” e il 35,6% il livello “molto”. Analogamente,

L'adeguatezza della dimensione partecipativa e collaborativa ha ottenuto una media pari a 4,315, mentre il contributo al miglioramento professionale ha riportato una media pari a 4,329. Quest'ultimo dato risulta particolarmente significativo, poiché evidenzia come il nuovo strumento non sia stato percepito esclusivamente come dispositivo di controllo, ma anche come possibile leva di sviluppo professionale. Dal punto di vista inferenziale, il test di Wilcoxon ha confermato che tutti gli item risultano significativamente superiori al valore neutro della scala, identificato nel punteggio 3, corrispondente ad "abbastanza". I valori di significatività sono risultati tutti inferiori a 0,001. Anche lo score composito ha mostrato una percezione significativamente superiore al livello intermedio della scala, con $W=2604,5$ e $p<0,001$. Questo risultato rafforza l'interpretazione secondo cui l'accettabilità dello strumento non deriva da una distribuzione casuale delle risposte, ma da un orientamento complessivamente favorevole e statisticamente evidente. Un elemento di particolare rilevanza riguarda l'affidabilità interna del questionario. Il coefficiente α di Cronbach è risultato pari a 0,900, valore indicativo di eccellente consistenza interna. Le correlazioni item-totale corrette sono risultate comprese tra 0,684 e 0,826, e l'eliminazione di ciascun item non ha determinato incrementi sostanziali dell' α . Ciò suggerisce che tutti gli item contribuiscono in modo coerente alla misurazione del costrutto complessivo di percezione e accettabilità del nuovo sistema valutativo. Anche l'analisi fattoriale esplorativa supporta la solidità preliminare dello strumento. L'indice KMO è risultato pari a 0,841, indicativo di buona adeguatezza campionaria, mentre il test di sfericità di Bartlett è risultato significativo, con $\chi^2=226,33$, $gdl=10$ e $p<0,001$. Il primo fattore ha mostrato un autovalore pari a 3,59 ed è risultato responsabile del 71,9% della varianza totale. I carichi fattoriali sono risultati elevati per tutti gli item, compresi tra 0,778 e 0,906. Tale configurazione suggerisce una struttura sostanzialmente monodimensionale del questionario, coerente con l'ipotesi che i cinque item misurino un unico costrutto latente riferibile alla percezione complessiva del nuovo sistema di valutazione. Un ulteriore aspetto discusso riguarda la familiarità pregressa dei valutatori con i modelli teorici utilizzati nella costruzione dello strumento. La familiarità completa con la Piramide di Miller è stata dichiarata dal 30,1% dei rispondenti, quella parziale dal 28,8%, mentre il 41,1% ha dichiarato di non conoscerla. Per il sistema AGENAS, invece, la familiarità completa è risultata pari al 49,3%, quella parziale al 20,5% e l'assenza di familiarità al 30,1%. La familiarità con i due modelli è risultata

significativamente associata, con $\chi^2=23,52$, $gdl=4$ e $p<0,001$, suggerendo una convergenza tra conoscenza del framework teorico e conoscenza del sistema istituzionale di riferimento. Tuttavia, le analisi bivariate non hanno evidenziato differenze statisticamente significative nello score composito in funzione della familiarità con i due modelli. Per la Piramide di Miller, il test di Kruskal-Wallis ha restituito $H=3,77$ e $p=0,152$; per il sistema AGENAS, $H=2,32$ e $p=0,313$. Anche le correlazioni ordinali sono risultate deboli e non significative, con $\rho=0,227$ per Miller e $p=0,054$, e $\rho=0,180$ per AGENAS e $p=0,128$. Questo dato può essere interpretato come un elemento favorevole alla comprensibilità dello strumento: la percezione positiva non sembra dipendere in modo determinante dalla conoscenza preliminare dei modelli teorici, ma appare diffusa anche tra i soggetti meno familiari con essi. La regressione lineare multipla esplorativa ha confermato tale andamento. Né la familiarità con la Piramide di Miller né quella con il sistema AGENAS sono risultate predittori significativi dello score composito. Il modello ha spiegato una quota limitata della variabilità complessiva, con $R^2=0,035$ e R^2 corretto=0,007. Questo risultato suggerisce che l'accettabilità dello strumento sia ampiamente distribuita tra i partecipanti e non vincolata alla conoscenza pregressa dei riferimenti teorico-istituzionali adottati. L'analisi qualitativa della domanda aperta ha ulteriormente arricchito l'interpretazione dei dati quantitativi. La categoria più rappresentata è stata l'assenza di suggerimenti o la conferma del modello, rilevata in 57 risposte, pari al 78,082% dei rispondenti. Una valutazione esplicitamente positiva o di adeguatezza percepita è emersa in 16 risposte, pari al 21,918%. Le aree di miglioramento hanno riguardato soprattutto l'affinamento di criteri, metodologia e indicatori, segnalato da 9 rispondenti, pari al 12,329%, seguito dal tema del confronto, feedback e partecipazione del valutato, emerso in 7 risposte, pari al 9,589%. Ulteriori osservazioni hanno riguardato monitoraggio longitudinale, formazione e miglioramento continuo, con 5 risposte pari al 6,849%, mentre il benchmarking esterno è stato citato da 1 rispondente, pari all'1,370%. I risultati evidenziano un'elevata accettabilità del nuovo sistema di valutazione delle performance GAAR. La marcata prevalenza di risposte positive agli item Q1–Q5 suggerisce che lo strumento sia percepito come chiaro, utile, comprensibile, pertinente rispetto alla dimensione collaborativa e potenzialmente orientato al miglioramento professionale. Tale risultato appare coerente con la finalità del modello, che non si limita alla misurazione della performance individuale, ma intende configurarsi

come dispositivo di governance orientato alla trasparenza, alla tracciabilità e al miglioramento continuo del processo valutativo. Dal punto di vista psicometrico preliminare, l'elevato α di Cronbach e la struttura monodimensionale emersa dall'analisi fattoriale esplorativa supportano l'utilizzo di uno score composito sintetico di percezione/accettabilità. Questo aspetto è particolarmente rilevante per l'eventuale prosecuzione del percorso di validazione dello strumento, poiché consente di disporre di un indicatore aggregato facilmente interpretabile e utilizzabile in successive analisi comparative o longitudinali. Le analisi bivariate e multivariate non hanno evidenziato un'associazione statisticamente significativa tra familiarità pregressa con i modelli di riferimento e percezione complessiva dello strumento. Tale risultato può essere interpretato in due modi complementari: da un lato, il nuovo sistema appare accettabile anche per i soggetti meno familiari con la Piramide di Miller o con il sistema AGENAS; dall'altro, la distribuzione fortemente positiva dello score riduce la variabilità disponibile per identificare predittori statisticamente significativi. In termini metodologici, questo elemento suggerisce l'opportunità di confermare i risultati su campioni più ampi e con disegni longitudinali, includendo ulteriori variabili professionali e organizzative.

IMPLICAZIONI PER IL SISTEMA DI AUTORIZZAZIONE E ACCREDITAMENTO

I risultati del presente studio assumono particolare rilevanza nel contesto del sistema di autorizzazione e accreditamento sanitario, in quanto evidenziano la possibilità di introdurre strumenti valutativi più strutturati, oggettivi e coerenti con i principi di qualità e accountability. La valutazione dei valutatori rappresenta infatti un elemento strategico per garantire l'affidabilità dell'intero sistema di accreditamento. In questo senso, l'introduzione di uno strumento basato su modelli teorici validati e integrato con i sistemi istituzionali di performance contribuisce a rafforzare la credibilità del processo valutativo e la fiducia nei confronti delle istituzioni sanitarie.

PUNTI DI FORZA DELLO STUDIO

Lo studio presenta diversi elementi di forza che ne rafforzano la solidità metodologica e l'interesse applicativo. In primo luogo, la progettazione del nuovo strumento valutativo è stata sviluppata attraverso un'integrazione sistematica tra le evidenze emerse dalla revisione della letteratura scientifica e le esigenze operative del contesto regionale di

riferimento, consentendo di coniugare rigore metodologico e concreta applicabilità. Un ulteriore punto di forza è rappresentato dall'adozione della Piramide di Miller quale framework teorico per la valutazione delle competenze professionali, modello ampiamente riconosciuto e validato in letteratura, che ha fornito una solida base concettuale per la strutturazione della scheda. La successiva integrazione con il Sistema di Misurazione e Valutazione della Performance di AGENAS ha inoltre permesso di sviluppare uno strumento coerente con gli indirizzi istituzionali nazionali e con i principi di accountability, trasparenza e miglioramento continuo propri del sistema sanitario pubblico. Infine, la sperimentazione diretta della scheda sull'intera popolazione dei valutatori GAAR attivi nella Regione Marche ha consentito di verificarne la fattibilità operativa e di acquisire dati preliminari sulla sua accettabilità, rappresentando un importante elemento di forza ai fini della futura implementazione del modello in contesti istituzionali analoghi.

LIMITI METODOLOGICI DELLO STUDIO

L'analisi presenta alcuni limiti. In primo luogo, il disegno trasversale consente di descrivere la percezione dei valutatori in un momento specifico, ma non permette inferenze causali sull'effetto del nuovo sistema di valutazione. In secondo luogo, la partecipazione volontaria può aver introdotto un bias di autoselezione, con possibile sovrarappresentazione di soggetti maggiormente motivati o favorevoli allo strumento. In terzo luogo, la distribuzione dei punteggi mostra un effetto ceiling, che può ridurre la capacità dei test di individuare differenze tra sottogruppi o associazioni con predittori esterni. Infine, il dataset disponibile non include variabili socio-professionali dettagliate, quali ruolo, anzianità nel GAAR, numero di audit effettuati o precedente esperienza valutativa, che avrebbero consentito analisi multivariate più articolate.

PROSPETTIVE DI SVILUPPO

I risultati ottenuti aprono la strada a possibili sviluppi futuri, tra cui:

- l'estensione dell'applicazione dello strumento ad altri contesti regionali;
- la validazione su campioni più ampi;
- l'integrazione con sistemi digitali di raccolta e analisi dei dati;

- l'eventuale personalizzazione del sistema valutativo in funzione dei diversi ruoli professionali.

CONCLUSIONI

Il presente studio ha consentito di progettare, sviluppare e sottoporre a una prima fase di validazione metodologica un nuovo sistema di valutazione delle performance dei valutatori del Gruppo di Autorizzazione e Accreditamento Regionale (GAAR), fondato sull'integrazione tra la **Piramide di Miller** e il **Sistema di Misurazione e Valutazione della Performance di AGENAS**. Tale integrazione ha permesso di coniugare un modello teorico consolidato per la valutazione delle competenze professionali con un sistema istituzionale orientato alla misurazione della performance, dando origine a uno strumento innovativo, coerente con il contesto organizzativo di riferimento e potenzialmente trasferibile ad altri ambiti della governance sanitaria. L'intero percorso di costruzione della scheda è stato sviluppato attraverso un processo metodologicamente strutturato, comprendente la revisione della letteratura scientifica, il coinvolgimento di esperti mediante focus group, la validazione preliminare istituzionale e la successiva implementazione sperimentale nell'ambito delle attività dell'Organismo Tecnicamente Accreditante della Regione Marche. Questo iter ha consentito di realizzare uno strumento fondato su solide basi teoriche e strettamente aderente alle esigenze operative del contesto applicativo, perseguendo principi di trasparenza, standardizzazione, oggettività e riproducibilità del processo valutativo. Le analisi psicometriche condotte nel corso della sperimentazione hanno fornito evidenze preliminari particolarmente favorevoli a sostegno della qualità metodologica dello strumento. In particolare, il **coefficiente α di Cronbach pari a 0,900** documenta un'eccellente consistenza interna della scheda, indicando un elevato livello di coerenza tra gli item che la compongono e confermando la capacità dello strumento di misurare in modo omogeneo il costruito oggetto di valutazione. Parallelamente, la verifica dei presupposti per l'analisi fattoriale, attraverso il **coefficiente Kaiser-Meyer-Olkin (KMO)** e il **test di sfericità di Bartlett**, ha confermato l'adeguatezza della matrice dei dati all'esplorazione della struttura latente dello strumento, mentre l'analisi fattoriale esplorativa ha evidenziato una configurazione sostanzialmente **monodimensionale**, con il primo fattore in grado di spiegare **il 71,9% della varianza complessiva**, rafforzando ulteriormente la validità di costruito della scheda proposta. Anche i risultati derivanti dalle analisi inferenziali contribuiscono a sostenere la solidità del modello sviluppato. L'assenza di associazioni statisticamente significative tra il livello di familiarità pregressa con la Piramide di Miller o con il Sistema

di Misurazione e Valutazione della Performance di AGENAS e la percezione complessiva dello strumento suggerisce infatti che il giudizio espresso dai valutatori non sia stato influenzato dalla conoscenza preliminare dei modelli teorici utilizzati nella sua progettazione, ma rifletta una valutazione sostanzialmente indipendente delle caratteristiche intrinseche della scheda e della sua applicabilità nel contesto operativo. Dal punto di vista scientifico, il lavoro rappresenta, per quanto a conoscenza degli autori, uno dei primi tentativi di sviluppare uno strumento specificamente orientato alla valutazione delle performance dei valutatori GAAR attraverso un approccio competence-based integrato con i sistemi istituzionali di performance management. Tale elemento conferisce al modello un carattere di originalità metodologica e apre interessanti prospettive di sviluppo nell'ambito della valutazione delle competenze professionali applicata ai processi di autorizzazione e accreditamento sanitario. Sotto il profilo organizzativo, l'adozione di una scheda strutturata, standardizzata e condivisa potrebbe contribuire a rendere maggiormente uniforme il processo di valutazione dei valutatori regionali, riducendo la variabilità interpretativa, aumentando la trasparenza dei criteri utilizzati e promuovendo una cultura orientata al miglioramento continuo della qualità. In tale prospettiva, lo strumento non si configura esclusivamente come un mezzo di misurazione delle performance, ma anche come un supporto allo sviluppo professionale, alla responsabilizzazione dei valutatori e al rafforzamento dell'intero sistema di governance dell'accreditamento sanitario. Pur in presenza di risultati ampiamente incoraggianti, il presente studio rappresenta una **fase preliminare del percorso di validazione** della scheda "Miller-AGENAS". Saranno pertanto necessari ulteriori studi prospettici e multicentrici, condotti su campioni più estesi e in differenti contesti organizzativi, finalizzati a confermare la stabilità temporale dello strumento, approfondirne ulteriormente le proprietà psicometriche, verificarne la validità esterna e valutarne la capacità discriminante rispetto a differenti livelli di performance professionale. In conclusione, le evidenze emerse dal presente lavoro consentono di affermare che la scheda di valutazione "**Miller-AGENAS**" possiede **solide preliminari caratteristiche di affidabilità interna, validità di costrutto e applicabilità operativa**, rappresentando un modello metodologicamente robusto e scientificamente fondato per la valutazione delle performance dei valutatori GAAR. La sua implementazione costituisce un importante passo verso la definizione di sistemi valutativi maggiormente

standardizzati, trasparenti e orientati allo sviluppo delle competenze, ponendo le basi per future applicazioni e per il consolidamento di un approccio evidence-based alla valutazione dei professionisti coinvolti nei processi di autorizzazione e accreditamento sanitario.

Bibliografia

- [1] *DGR 1982 del 18 dicembre 2023 "Modifica DGR n. 522 del 23/4/2018 L.R. 30 settembre 2016, n. 21 e s.m.i. Costituzione e funzionamento dell'Organismo Tecnicamente Accreditante (OTA)".*
- [2] *Decreto del Direttore dell'Agenzia Regionale Sanitaria N.73 del 09/07/2024: Costituzione e funzionamento del Gruppo di Autorizzazione ed Accredimento Regionale (GAAR) Revisione del decreto del Settore del Territorio e integr. Socio Sanitaria n.28 17/07/17.*
- [3] *Decreto del Dirigente del settore territorio ed integrazione socio sanitaria -ARS. N.31 del 17 Luglio..*
- [4] *D. C. M. M. G. Marini, Accountability e performance del Servizio Sanitario Nazionale italiano: misurazione, valutazione e rendicontazione, «I paper dell'Osservatorio», Osservatorio AIR. P 1/2025, Febbraio 2025..*
- [5] *A. Donabedian, The quality of care. How can it be assessed?, JAMA, 1988..*
- [6] *J. Denis, Accountability in healthcare organizations and systems, Healthcare Policy, 2014..*
- [7] *G. Marini, D. Chiarito, M. Mele, op. cit..*
- [8] *Agenas, Agenzia Sanitaria per i Servizi Sanitari Regionali. "Sistema di Misurazione e Valutazione della Performance. Art-7, comma1. D.Lgs. 150/2009. Aggiornamento 2024., 2024..*
- [9] *Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. Zachary Munn, Micah D. J. Peters, Cindy Stern, Catalin Tufanaru, Alexa McArthur and Edoardo Aromataris.*

- [10] *The PICO strategy for the research question construction and evidence search*
Cristina Mamédio da Costa Santos I, Cibeles Andrucioli de Mattos Pimenta,
Moacyr Roberto Cuce Nobre.
- [11] *Successful combination of database search and snowballing for identification of
primary studies in systematic literature studies*. Author links open overlay
panel Claes Wohlin, Marcos Kalinowski, Katia Romero Felizardo, Emilia Mendes.
- [12] AGENAS, *Sistemi di Misurazione e Valutazione delle Performance*. Art. 7,
comma 1, D. Lgs. 150/2009, Aggiornamento 2024.
- [13] n. 1. Decreto Legislativo 27 ottobre 2009, *Attuazione della legge 4 marzo 2009, n.
15, in materia di ottimizzazione della produttività del lavoro pubblico e di
efficienza e trasparenza delle pubbliche amministrazioni.*, 11/2009.
- [14] Miller, G. E. (1990). *The assessment of clinical skills/competence/performance*.
Academic Medicine, 65(9), S63...
- [15] Rhind, S. M., MacKay, J., Brown, A. J., Mosley, C. J., Ryan, J. M., Hughes, K. J.,
& Boyd, S. (2021). *Developing Miller's Pyramid to Support Students' Assessment
Literacy*. *Journal of Veterinary Medical Education*, 48(2), 158–162...
- [16] Cruess, R. L., Cruess, S. R., & Steinert, Y. (2016). *Amending Miller's Pyramid to
Include Professional Identity Formation*. *Academic Medicine*, 91(2), 180..
- [17] D. C. M. M. G. Marini, *Accountability e performance del Servizio Sanitario
Nazionale italiano: misurazione, valutazione e rendicontazione*, «I paper
dell'Osservatorio», Osservatorio AIR. P 1/2025, Febbraio 2025.
- [18] Domínguez-Torres, L. C., & Vega-Peña, N. V. (2023). *Pyramids in medical
education: Their conceptualization and utility summarized*. *Revista Colombiana
De Obstetricia Y Ginecología*, 74(2), 163–174.
<https://doi.org/10.18597/rcog.3994>.

[19] *DGR 1982 del 18 dicembre 2023 "Modifica DGR n. 522 del 23/4/2018 L.R. 30 settembre 2016, n. 21 e s.m.i. Costituzione e funzionamento dell'Organismo Tecnicamente Accreditante (OTA)"..*

Allegati

ALLEGATO 1

Scheda di valutazione individuale GAAR

Anno 2024

Nome

Cognome

Items di valutazione:

1. EFFETTUAZIONE DELLE VISITE ISPETTIVE COORDINATE DALL'OTA

Il fattore valuta la disponibilità all'esecuzione delle visite nelle strutture sanitarie e socio-sanitarie

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

2. ACCURATEZZA NELLA REDAZIONE DEI VERBALI DI VISITA ISPETTIVA

Il fattore valuta l'impegno e l'attenzione nella stesura del verbale e dell'invio all'OTA entro massimo 5 giorni dalla visita

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

3. ESECUZIONE DEL PRE-AUDIT ED ACCURATEZZA NELLA COMPILAZIONE DEL FORMAT "PIANO DI AUDIT"

Il fattore valuta la sistematica partecipazione alle attività di pre-audit con particolare riferimento alla compilazione della scheda "piano di audit" in funzione di team leader o componente del gruppo incaricato

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

4. ACCURATEZZA NELLA REDAZIONE DEI VERBALI DI VISITA PER AUTORIZZAZIONE E ACCREDITAMENTO

Il fattore valuta l'impegno e l'attenzione nella stesura del verbale e dell'invio all'OTA entro massimo 10 giorni dalla visita

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

5. ACCURATEZZA NELLA VALUTAZIONE DI ASSOLVIMENTO PRESCRIZIONI RILASCIATE ALLA STRUTTURA SOGGETTA A VISITA PER AUTORIZZAZIONE O ACCREDITAMENTO

Il fattore valuta l'impegno e l'attenzione nella valutazione della conformità delle evidenze documentali in caso di prescrizioni rilasciate in sede di visita di autorizzazione o accreditamento: stesura del verbale e invio all'OTA entro massimo di 10 giorni

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

6. ACCURATEZZA NELLA VALUAZIONE DELLE EVIDENZE DOCUMENTALI DI ACCREDITAMENTO

Il fattore valuta l'impegno e l'attenzione nella valutazione delle evidenze documentali per l'accreditamento avanzato e/o di eccellenza: stesura del verbale e invio all'OTA entro massimo 15 giorni

1	2	3	4	5	6	7	8	9	X
---	---	---	---	---	---	---	---	---	---

7. PARTECIPAZIONE AGLI EVENTI FORMATIVI SPECIFICI PER IL GAAR PROGRAMMATI DALL'OTA

Il fattore valuta l'impegno e la disponibilità ad aggiornarsi, ad arricchire le proprie conoscenze professionali nonché ad applicare e diffondere quanto appreso

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

8. PARTECIPAZIONE ALLE RIUNIONI INDETTE DALL'OTA

Il fattore valuta la disponibilità attiva alle riunioni in presenza o a distanza

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

9. DISPONIBILITÀ A COLLABORARE NELLA INDIVIDUAZIONE DI SOLUZIONI ORGANIZZATIVE INNOVATIVE

Il fattore valuta la dimostrata propensione nell'individuare soluzioni organizzative innovative anche in caso di criticità riscontrate in corso di visita

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Range valori singolo item: eccellente (10); buono (8 - 9); sufficiente (6 - 7); insufficiente (1-5)

Legenda – correlazione punteggio ottenuto e giudizio di valutazione finale:

- eccellente (80 – 90)
- buono (66 – 79)
- sufficiente (54 – 65)
- insufficiente (≤ 53)

Nota: nel caso in cui non siano state assegnate dall'OTA, attività di "visita ispettiva" per gli items di valutazione n. 1 e n. 2, agli stessi verrà attribuito il punteggio massimo.

Giudizio di valutazione finale: (scrivere il punteggio ottenuto)

Punteggio totale ottenuto	
---------------------------	--

Eventuali azioni di miglioramento concordate

CRITICITÀ EMERSE	AZIONI DI MIGLIORAMENTO

Il Valutato concorda con il giudizio formulato dal Valutatore?

SI	NO
----	----

Data _____ Firma valutato _____

Data _____ Firma Coordinatore OTA _____

Annotazioni e commenti in caso di difformità di giudizio e/o non accettazione: _____ _____ _____
--

Data _____ Firma valutato _____

Data _____ Firma Coordinatore OTA _____

ALLEGATO 2

Scheda di valutazione team di valutazione GAAR da parte delle strutture

Struttura visitata: _____ Data: __/__/__

Team Leader	
Valutatore	
Valutatore	
Valutatore	
Valutatore	
Valutatore	
Valutatore	

Items di valutazione:

parte generale

puntualità del team all'audit

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Note _____

rispetto del programma cronologico e di attività prefissato

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Note _____

capacità di risoluzione situazioni critiche che avrebbero inficiato l'audit

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Note _____

competenza specifica dal gruppo di valutatori estrinsecata in corso di audit

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Note _____

coesione operativa del gruppo di valutatori apprezzata in corso di audit

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Note _____

Competenza e riservatezza dei valutatori rispetto a soggetti terzi

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Note _____

capacità di condivisione delle prescrizioni proposte con i valutati

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Note _____

Team leader

- Capacità di creare un clima collaborativo
- Puntualità
- Capacità di coordinamento del gruppo (percezione della partecipazione attiva e propositiva al processo/percorso da parte di tutti i componenti del team di valutatori, autorevolezza)
- Rispetto del programma concordato
- Riservatezza

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Note _____

Valutatori

- Capacità di creare un clima collaborativo
- Puntualità
- partecipazione attiva e propositiva al processo/percorso
- Rispetto del programma concordato per la parte assegnata
- riservatezza

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Note _____

Apporre la valutazione contrassegnando una casella per ogni singolo ITEM con il seguente criterio:

1 minimamente soddisfatto

10 altamente soddisfatto

Note _____

Il Legale Rappresentante (nome /qualifica/incarico)

ALLEGATO 3**ORGANISMO TECNICAMENTE ACCREDITANTE
Valutazione delle Performance GAAR**

Anno di riferimento: 2025

VALUTATORE GAAR**Nome:****Cognome:***(Modelli concettuale di riferimento: Piramide di Miller e Sistema di Valutazione AGENAS)***A. Performance Operativa in Fase di Pre-Audit**

Item	Standard atteso	Criteri osservabili	Punteggio			
Esecuzione del pre-audit	Pianificazione dell'incontro preliminare (pre-audit) con tempistiche adeguate rispetto alla data della visita di audit	a) Rispetto dei tempi previsti di programmazione del pre-audit b) Consegna del verbale di pre-audit all'OTA nelle modalità e nei tempi previsti	2	4	6	8
Accuratezza nella redazione dei verbali di pre-audit	Verbali di pre-audit completi, coerenti e conformi ai requisiti metodologici	a) Completezza dei contenuti b) Coerenza e aderenza alla metodologia previste c) Accuratezza formale e tecnica d) Tracciabilità delle evidenze	2	4	6	8
Accuratezza nella valutazione delle evidenze presentate dalla struttura	Il valutatore analizza in modo accurato, sistematico e metodologicamente corretto tutte le evidenze presentate dalla struttura, verificandone pertinenza, attendibilità, coerenza con gli standard previsti e tracciabilità. La valutazione è imparziale, completa e supportata da elementi oggettivi.	a) Chiarezza nella registrazione delle evidenze b) Oggettività e neutralità c) Accuratezza metodologica d) Completezza dell'analisi	2	4	6	8
Punteggio parziale Area valutativa A) [punteggio max. attribuibile 24]						

B. Performance Operativa in Fase di Audit

Item	Standard atteso	Criteri osservabili	Punteggio			
Accuratezza nella redazione dei verbali di audit	Verballi di audit completi, coerenti e conformi ai requisiti metodologici	a) Completezza dei contenuti b) Coerenza e aderenza alla metodologia previste c) Accuratezza formale e tecnica d) Tracciabilità delle evidenze	2	4	6	8
Accuratezza nella valutazione delle evidenze presentate dalla struttura	Il valutatore analizza in modo accurato, sistematico e metodologicamente corretto tutte le evidenze presentate dalla struttura, verificandone pertinenza, attendibilità, coerenza con gli standard previsti e tracciabilità. La valutazione è imparziale, completa e supportata da elementi oggettivi.	a) Chiarezza nella registrazione delle evidenze b) Oggettività e neutralità c) Accuratezza metodologica d) Completezza dell'analisi	2	4	6	8
Coerenza tra evidenze raccolte e valutazioni formulate	Il valutatore formula giudizi e valutazioni pienamente coerenti con le evidenze raccolte durante l'audit, assicurando corrispondenza tra ciò che osserva/documenta e quanto conclude, senza discrepanze, sovra- o sotto-valutazioni.	a) Tracciabilità tra evidenze e giudizi. b) Assenza di discrepanze o valutazioni non supportate. c) Classificazione conforme agli standard GAAR.	2	4	6	8
Punteggio parziale Area valutativa B) [punteggio max. attribuibile 24]						

C. Conclusione visita di Audit

Item	Standard atteso	Criteri osservabili	Punteggio			
Completezza griglia prescrizioni	Prescrizioni formulate correttamente	Coerenza con non conformità, dettaglio	2	4	6	8
Completezza scheda fatturazione	Corretta compilazione secondo criteri OTA	Inserimento dei dati previsti	2	4	6	8
Rispetto della tempistica di consegna del verbale	Il verbale è stato consegnato alla struttura e all'OTA entro i tempi prestabili	a) Tempo di consegna del verbale di audit alla struttura b) Tempo di consegna del verbale all'OTA	2	4	6	8
Punteggio parziale Area valutativa C) [punteggio max 24]						

D. Partecipazione e collaborazione istituzionale

Item	Standard atteso	Criteri osservabili	Punteggio			
Partecipazione eventi formativi GAAR	Partecipazione attiva e continuativa	Presenza, puntualità, applicazione contenuti	2	4	6	8
Partecipazione riunioni OTA	Contributo attivo	Partecipazione, domande, proposte	2	4	6	8
Disponibilità partecipazione alle visite	Affidabilità e rispetto convocazioni	Risposte tempestive, rispetto scadenze	2	4	6	8
Punteggio parziale Area valutativa D) [punteggio max 24]						

Livelli di giudizio complessivi (sistema AGENAS + Miller)

- 0–24 → Non idoneo
- 25–48 → Parzialmente idoneo (Knows)
- 49–72 → Idoneo ((Knows How/Shows How)
- 73–96 → Eccellente (Does)

Giudizio di valutazione finale: (scrivere il punteggio ottenuto)

Punteggio totale ottenuto	
---------------------------	--

Eventuali azioni di miglioramento concordate

CRITICITÀ EMERSE	AZIONI DI MIGLIORAMENTO

Il Valutato concorda con il giudizio formulato dal Valutatore?

☐SI ☐NO

Data _____ Firma valutato _____

Data _____ Firma Coordinatore OTA _____

Annotazioni e commenti in caso di difformità di giudizio e/o non accettazione: _____ _____ _____ _____

Data _____ Firma valutato _____

Data _____ Firma Coordinatore OTA _____

ALLEGATO 4

Questionario breve sulla percezione del nuovo sistema di valutazione GAAR

PERCEZIONE NUOVO SISTEMA DI VALUTAZIONE DELLE PERFORMANCE GAAR

Il presente questionario ha lo scopo di rilevare la tua percezione complessiva in merito al modello di valutazione delle performance GAAR recentemente introdotto, con particolare riferimento alla sua chiarezza, utilità e coerenza con le attività svolte. La rilevazione è condotta indipendentemente dall'esito della valutazione individuale conseguita, con l'obiettivo di raccogliere opinioni, livello di gradimento e suggerimenti utili al miglioramento del modello stesso. Le risposte fornite contribuiranno a un'analisi conoscitiva volta a supportare eventuali azioni di revisione e ottimizzazione del sistema di valutazione.

Le risposte sono anonime.

Scala di risposta (dove indicato):

1 = Per nulla

2 = Poco 3 = Abbastanza

4 = Molto

5 = Completamente

N.	Domanda	Modalità di risposta				
1	Nel complesso, quanto ritieni chiaro il nuovo sistema di valutazione delle Performance GAAR?	1	2	3	4	5
2	Ritieni utile articolare la valutazione delle performance in relazione alle diverse fasi operative dell'attività GAAR (pre-audit, audit, conclusione della visita)?	1	2	3	4	5
3	Quanto ritieni chiari i criteri di valutazione utilizzati nel nuovo modello di valutazione delle performance GAAR?	1	2	3	4	5
4	Quanto ritieni pertinente includere la Partecipazione e collaborazione istituzionale tra i criteri di valutazione delle performance GAAR?	1	2	3	4	5
5	Quanto ritieni che il sistema sia utile per favorire il miglioramento professionale del valutatore?	1	2	3	4	5
6	Prima dell'introduzione di questo sistema, conoscevi la Piramide di Miller come modello di valutazione delle competenze?	Sì ☐ No ☐				
7	Prima dell'introduzione di questo sistema, conoscevi il Sistema di Valutazione AGENAS?	Sì ☐ No ☐				

In una frase, hai un suggerimento per migliorare il modello di valutazione?

ALLEGATO 5

Risposte alla domanda aperta

N°	Risposta
1	se fosse possibile inserirei la codifica del modulo con revisione in relazione al sistema gestione qualità
2	Mi sembra una buona proposta
3	No lo trovo idoneo Grazie
4	--
5	differenziare i criteri in base ai ruoli assunti e ai contesti valutati
6	nessun suggerimento
7	Nessuno
8	Peserei un pochino di più la parte operativa dell' audit rispetto a quella di preaudit
9	Ritengo il modello adeguato senza necessità di modifiche
10	Nessuna
11	Non ho suggerimenti
12	Al mometo nessun suggerimento
13	Nessuna osservazione
14	potrebbe essere utile effettuare un colloquio , magari visto che siamo in tanti , procedere con un campione a rotazione annuale
15	
16	Nessun suggerimento
17	Assolutamente completa!
18	Al momento non ho nessuna proposta
19	La valutazione dovrebbe essere effettuata dopo un congruo periodo di applicazione ed analisi dei risultati...
20	Va bene così
21	Non ho suggerimenti in proposito.
22	nessun suggerimento se non quello di valutare nel corso dell'anno in caso di necessita di introdurre una valutazione intermedia
23	Ritengo che sia già completo il modello utilizzato

24	il modello è chiaro e le parti evidenziate come da migliorare sono state discusse...
25	NO
26	No, penso che il modello di valutazione introdotto così articolato sia chiaro e utile...
27	non ho particolari suggerimenti
28	No
29	Nessun suggerimento
30	/
31	ritengo il modello esaustivo in ogni sua parte
32	no
33	Sostengo da sempre la necessità di occasioni di confronto sia tra GAAR che tra GAAR e OTA
34	L'attuale modello di valutazione mi sembra esaustivo e dettagliato...
35	Per me è la prima volta...
36	Qualità delle modalità di audit
37	Miglioramento continuo sulla base dei report di valutazione...
38	no
39	Andrebbe meglio specificata la metodologia di calcolo...
40	credo che sia un modello di valutazione piuttosto completo...
41	nessuno
42	No
43	La piramide di miller è già un ottimo modello...
44	No
45	NO
46	Non ho suggerimenti o proposte per migliorare il modello di valutazione
47	nulla da segnalare
48	nessuna
49	Nessun suggerimento
50	Nessuna proposta, va bene così
51	no, lo trovo già molto buono
52	Una esperienza fuori Regione...
53	nessuna

54	nessuna
55	Non ho suggerimenti da dare
56	sempre interessante anche il confronto collegiale
57	nulla
58	Confrontare nel tempo il modello...
59	No non ho suggerimenti o proposte migliorative
60	Al momento non ho particolari suggerimenti...
61	Trovo il modello valido...
62	No
63	No
64	Al momento non ho nessuna proposta. Cordiali saluti
65	In questo momento direi di no
66	No
67	Sarebbe opportuno inserire anche l'esito della valutazione...
68	No
69	no
70	A partire dalla formazione organizzata...
71	no
72	No, è sufficientemente adeguato a mio avviso.
73	