

DOCUMENTO UMI SULLA VQR

L'Assemblea UMI tenutasi a Bologna l'11/10/2013 ha esaminato e discusso i risultati della recente VQR. Il presente documento sintetizza gli esiti di quella discussione.

1) L'UMI riconosce che la valutazione è necessaria al miglioramento della qualità della ricerca, ed ha visto con favore l'avvio di questi processi nel nostro paese. L'UMI sottolinea con piacere che il mondo accademico ha, nel suo complesso, condiviso questo positivo atteggiamento, accettando di buon grado di sottoporsi, unico nel settore pubblico, a un vaglio delle proprie attività

2) *Usi propri e impropri dei risultati della VQR.*

- La VQR ha lo scopo di fornire informazioni obiettive sull'attività scientifica delle strutture che servano innanzitutto agli addetti ai lavori per migliorarne le prestazioni.
- La VQR ha, tra i suoi obiettivi, quello di fornire strumenti per rendere più efficiente ed efficace la distribuzione di risorse umane e finanziarie alle varie strutture. **Si auspica pertanto che venga confermata la distribuzione, in base alle risultanze della VQR, di una quota premiale del FFO che, si spera, veda un necessario incremento per rendere questa quota cospicua.** La mancata assegnazione di tale quota vanificherebbe, in gran parte, lo sforzo fatto dal sistema università-ricerca con il sottoporsi alla VQR.
- La distribuzione di risorse dovrebbe, a regime, premiare sia chi ha ottenuto i risultati migliori in assoluto, **sia chi è migliorato** rispetto alla precedente valutazione, in modo da fungere da stimolo non solo al mantenimento dell'eccellenza ma anche al miglioramento della qualità.
- Scopo della VQR è anche quello di individuare le situazioni critiche, onde indirizzare risorse verso la loro risoluzione. È importante tuttavia che ci si assicuri che a gestire tali risorse non siano proprio coloro che hanno creato, o quanto meno non risolto, le situazioni critiche in questione.
- È opportuno esaminare la possibilità che parte delle risorse distribuite dal Ministero agli Atenei in base alla VQR sia vincolata a essere destinata alle sottostrutture (dipartimenti) che meglio hanno figurato nella valutazione.
- La VQR, nel dare un quadro obiettivo della situazione della ricerca pubblica nel paese, non dovrebbe partire da presupposti ideologici, quali, ad esempio, quello di considerare mediocre il 50% "inferiore" della produzione scientifica. In particolare, la VQR non dovrebbe avere come obiettivo la creazione di classifiche (*ranking*) di università, ma dovrebbe raccogliere dati che consentano di raggruppare le strutture in classi di merito in base alla loro efficienza e qualità (*rating*). In questa operazione è necessario tenere separate università di dimensioni diverse, e distinguere anche fra università generaliste e università specializzate.
- **In nessun modo, e a nessun titolo, i risultati della VQR vanno usati per la valutazione dei singoli.** Va perciò assicurata la massima riservatezza dei dati personali risultanti dalla VQR, proteggendoli da eventuali attacchi e pressioni esterne volte a venirne a conoscenza e, peggio ancora, a renderli pubblici o usarli per fini impropri. È utile ricordare che la raccolta dei dati per la VQR è stata pensata per essere applicata a campioni numerosi, e, quando applicata a campioni piccoli, i dati forniti perdono di senso, o danno luogo perfino a risultati distorti in quanto non rappresentativi della ricerca dei singoli.

3) *Stabilità del sistema e dei meccanismi di valutazione.*

- È essenziale che gli esercizi di valutazione si ripetano a intervalli regolari, coprendo archi di tempo abbastanza ampi da dare risultati significativi, ma non eccessivamente lunghi, in modo da evitare che siano eventi isolati, non correlati l'uno con l'altro.
- Uno dei pregi di un sistema stabile e periodico di valutazione è la possibilità di seguire l'evoluzione della qualità delle strutture nel tempo: è quindi di primaria importanza elaborare meccanismi che permettano di confrontare i risultati di una VQR con quelli della successiva, ed è auspicabile che vengano usate metodologie omogenee nelle varie valutazioni periodiche.
- È necessario dunque non solo stabilire, ma **mantenere** una cadenza per la VQR. Sarebbe bene, ad

esempio, sapere già fin d'ora con sicurezza quando verrà effettuata la prossima.

- È necessario che i criteri e i meccanismi che saranno usati per la prossima VQR siano resi noti con congruo anticipo, in modo che la comunità abbia una chiara idea di quali siano i comportamenti virtuosi da porre in essere per ben riuscire. Inoltre è indispensabile che i criteri e meccanismi scelti siano tali da non indurre la comunità ad assumere comportamenti opportunistici e antiscientifici al solo scopo di migliorare artatamente la valutazione senza incidere su un aumento della qualità della ricerca (ad esempio, comportamenti tesi ad aumentare artificialmente il numero di citazioni).

4) *Formazione e direzione dei GEV e scelta dei valutatori.*

- Senza nulla togliere alla qualità e competenza dei componenti dei GEV della VQR 2004-2010, è opportuno che in futuro le procedure di designazione siano più tempestive e limpide. Muoversi con chiarezza di intenti e per tempo è indispensabile per evitare decisioni affrettate che rischiano di apparire poco trasparenti;
- è necessario che il MIUR predisponga al più presto, anche sulla base dell'esperienza della passata VQR, una lista pubblica di possibili valutatori cui ci si possa iscrivere rispondendo a criteri che considerino le esperienze già fatte nell'ambito della valutazione, la qualità scientifica e le competenze riconosciute a livello nazionale ed internazionale (riferendosi sia a indicatori bibliometrici che non bibliometrici) e tenendo presente le specificità delle varie discipline;
- per la individuazione dei componenti dei GEV, va prevista una chiamata pubblica di disponibilità mirata allo specifico esercizio di valutazione che si va ad effettuare. Tale chiamata deve essere tempestiva (va fatta almeno un anno prima della formazione dei GEV e va tenuta aperta per un periodo ragionevolmente lungo) e deve essere riservata a coloro che abbiano specifici requisiti di alta qualità scientifica ed esperienza di valutazione;
- è opportuno che anche per la scelta dei presidenti dei GEV venga aperta una tempestiva chiamata pubblica riservata a coloro che abbiano specifici requisiti di altissima qualità scientifica e notevole esperienza di valutazione;
- per la formazione dei GEV, e soprattutto per la scelta dei loro presidenti, e in ogni caso per la scelta di referenti e consulenti nel mondo accademico e della ricerca, sarebbe bene prevedere una qualche forma di consultazione delle comunità scientifiche, volta non alla difesa di interessi corporativi, ma alla evidenziazione di esigenze peculiari e qualità specifiche delle singole discipline.

5) *Valutazione dell'Istituto nazionale di Alta Matematica 'F. Severi'*

- L'INdAM ha una caratteristica peculiare tra gli Enti di Ricerca vigilati dal MIUR: è privo di un proprio personale di ricerca. Infatti l'INdAM svolge principalmente un ruolo di agenzia di finanziamento, fornendo supporto ai docenti universitari o dipendenti di altri enti che siano aderenti ai suoi gruppi di ricerca, e finanziando inoltre iniziative di interesse generale, quali borse di studio, assegni di ricerca, incontri scientifici, ecc. Queste iniziative sono possibili solo grazie al contributo **volontario e gratuito** di esponenti della comunità matematica italiana. I meccanismi di valutazione dell'INdAM devono dunque tenere conto di queste peculiarità. Attualmente invece essi finiscono per penalizzare la comunità matematica, con la richiesta, per un congruo numero di suoi membri, di un maggior numero (sei invece di tre per la passata VQR) di prodotti di ricerca di elevata qualità rispetto a quello richiesto ad altre comunità scientifiche. Tale richiesta ha posto peraltro problemi non banali circa la proprietà intellettuale dei prodotti: è essa dell'INdAM o delle strutture di appartenenza?
- Per il futuro una **proposta** è questa: si preveda, nel caso di un numero maggiore della norma di prodotti che risultino di elevata qualità, dei meccanismi che premiano tale esito sia per l'INdAM che per le strutture di appartenenza. Ciò potrebbe volgere l'attuale situazione di svantaggio in una sfida vantaggiosa per tutta la comunità a produrre di più e meglio.

6) *Classificazione delle riviste, peer review e bibliometria.*

- La struttura della ricerca matematica ha alcune importanti peculiarità, tra cui: tempi lunghi fra la presentazione di un articolo ad una rivista e la sua pubblicazione, tipicamente variabili fra 6 mesi e 2 anni; vita media di un articolo misurata in anni, spesso in decenni; scarsa correlazione fra difficoltà e importanza di un argomento di ricerca e le dimensioni della comunità che se ne occupa; ecc. Tutto ciò porta a ritenere, opinione abbastanza diffusa tra i matematici, che in valutazioni di tipo bibliometrico sia più importante la qualità della sede di pubblicazione di un articolo, piuttosto che il numero di citazioni ricevute, soprattutto se queste sono riferite ai primi anni dopo la pubblicazione. È dunque opportuno, in ambito matematico, non dare una netta prevalenza al numero di citazioni ma favorire la compresenza di più metodi di valutazione (citazioni, importanza della sede di valutazione, peer review, altro), che possano svolgere un ruolo di controllo reciproco ed essere coniugati in modo diverso a seconda delle caratteristiche di ciò che si deve valutare e del momento in cui avviene la valutazione.
- La qualità della sede di pubblicazione non è indipendente dagli argomenti di ricerca. Gli esempi più evidenti riguardano matematica applicata, didattica e storia della matematica. Per loro natura queste discipline hanno modalità di pubblicazione e un pubblico alquanto diversi da quelli del resto della matematica. Quindi, mentre i criteri di valutazione della sede di pubblicazione possono essere alquanto omogenei per settori come analisi matematica, calcolo delle probabilità, geometria, algebra, fisica matematica, ecc. è invece indispensabile usare criteri ad hoc per le su citate discipline. Ad esempio, la storia della matematica (ma analoghe considerazioni valgono per la didattica), pur con una sua specificità che la lega fortemente all'ambito matematico, condivide molti aspetti della sua produzione con la storia della scienza, disciplina che fa capo a settori umanistici, per i quali si usano giustamente criteri ben diversi da quelli delle cosiddette "scienze dure". Quanto alla matematica applicata è giusto continuare ad indirizzare pubblicazioni rivolte prevalentemente ad un settore diverso da quello della matematica a quest'ultimo piuttosto che a quello matematico, così come è necessario tenere conto di altre specificità in questo ambito, quali una diversa importanza, rispetto alla matematica pura, di atti di convegni.
- La qualità di una sede di pubblicazione non è costante nel tempo; inoltre, frequentemente alcune riviste cessano e nuove nascono. Quindi qualsiasi valutazione della qualità della sede di pubblicazione deve essere soggetta a revisioni e aggiornamenti periodici, non necessariamente legati in senso stretto ad una VQR. Un tale esercizio può essere infatti molto utile per elevare la qualità delle riviste e per contrastare il proliferare di riviste di livello basso accompagnato da una tendenza (talvolta accentuata proprio da equivoci sugli obiettivi della valutazione) ad aumentare la quantità della propria produzione, ma non la sua qualità. Riguardo al primo aspetto si osserva il recente sorgere di nuove riviste, varie "open access", dettato da evidenti motivazioni commerciali più che scientifiche. Il secondo fenomeno va fortemente scoraggiato, per cui mai i meccanismi di valutazione dovranno incoraggiare la quantità piuttosto che la qualità. Altro obiettivo cui puntare con un'attenta, periodica valutazione della qualità è quello della spinta al miglioramento delle riviste, specie le italiane, che avranno un buon motivo di competere per ottenere un buon livello di giudizio.
- Ma come effettuare un "ranking" appropriato delle riviste in gruppi di qualità? A valle della determinazione di quali siano le riviste "accettabili", la suddivisione in quattro livelli adoperata nella recente VQR è di per sé condivisibile. Tuttavia l'appartenenza a ciascuno dei quattro livelli dovrebbe essere accertata non solo in base a meri criteri numerici (tipo l'IF), ma anche a criteri di qualità che comprendano il livello scientifico dei "managing editors" e degli "editorial boards", dati storici sulla costanza nel tempo della qualità della rivista, la considerazione a livello internazionale, ecc. In questo ordine di idee andrebbe rivista la scelta di stabilire a priori in modo rigido la percentuale di riviste da destinare a ciascun livello: ad esempio, stabilendo, come si è fatto nella recente VQR, che nel quarto livello va messo il 50% delle riviste, si suggerisce, in sostanza, che il 50% della produzione scientifica del nostro paese è mediocre. Meglio forse fissare una percentuale massima di riviste che possano essere inserite in ciascun livello.
- Anche se per certi versi comprensibile, andrebbe in linea di massima evitato che una stessa rivista stia in livelli diversi per discipline diverse (e.g., analisi matematica e geometria, cosa accaduta nella recente VQR). Questo potrebbe accentuare il fenomeno, in parte già in atto, di specializzazione delle riviste a danno di quelle generaliste e interdisciplinari, che costituiscono un

terreno di confronto e collegamento fra scienziati che operano su terreni diversi. Quando una tale ripartizione fosse suggerita da meri criteri numerici, un giudizio competente dovrebbe assicurarsi se la cosa abbia effettivamente senso dal punto di vista scientifico.

- Per quanto riguarda le citazioni, e ben comprendendo che quanto andiamo a proporre non sia di facile realizzazione, la cosa migliore sarebbe di disporre di una molteplicità di dati, che comprendano numero totale di citazioni, numero di autocitazioni, numero di autori distinti che citano, ecc. e confrontarli tra loro per estrarne il maggior numero possibile di informazioni. In presenza di evidenti distorsioni (ad esempio: il numero di autocitazioni supera nettamente quello delle citazioni di altri autori; tutte le citazioni di un dato articolo, o autore provengono da un numero molto ristretto di altri autori, ecc.) il GEV potrebbe intervenire col peer review per risolvere eventuali dubbi.

7) *Meccanismi di valutazione.*

- La presenza di "inattivi" accertata dalla recente VQR merita un approfondimento. È infatti probabile che si tratti, in molti casi, di personale tutt'altro che inattivo, che fornisce servizi essenziali all'università operando, magari maggiormente rispetto alla media dei docenti, in campi quali la didattica, l'amministrazione, le attività di terza missione. Si ritiene quindi opportuno, per fornire un quadro completo della situazione, esaminare (o indurre le strutture ad esaminare) nello specifico le cause di inattività, accertando quali siano i maggiori servizi aggiuntivi resi dagli "inattivi" a vantaggio dei dipartimenti e degli atenei.
- Meccanismi di valutazione a distanza come quelli usati per la recente VQR non forniscono un'immagine completa della situazione dei dipartimenti e degli atenei per quanto riguarda gli aspetti della ricerca e delle attività di terza missione non finalizzati unicamente alla produzione di pubblicazioni. Si parla ad esempio di redazione di trattati o monografie di alto valore scientifico, organizzazione di convegni e edizione dei relativi atti, seminari, interventi sul territorio, ecc. Si suggerisce dunque un'acquisizione più estensiva di dati necessari per una più completa e articolata valutazione, incoraggiando anche meccanismi di autovalutazione.
- Un punto delicato della VQR 2004-2010 consiste nella rigida suddivisione a priori dei prodotti nelle categorie eccellenti/buoni/accettabili/limitati corrispondenti al cadere rispettivamente nel primo 20%/secondo 20%/terzo 10%/ultimo 50% della produzione mondiale. Questa suddivisione, come già accennato dianzi, presenta vari problemi. Infatti non pare realistico stabilire a priori che la metà della produzione scientifica non solo del nostro paese, ma del mondo intero, sia da considerarsi "limitata" e che pertanto **valga zero**. Pur volendo mantenere la suddivisione in quattro livelli, appare, come suggerito nel caso delle riviste, meglio stabilire la percentuale massima di prodotti che possano essere inseriti in ciascun livello. Appare anche difficilmente comprensibile la scarsa ampiezza del terzo livello, quello dei prodotti definiti "accettabili", rispetto agli altri. A tale proposito, giova mettere in risalto un vero e proprio problema linguistico: nell'uso corrente e nell'esperienza di ciascuno di noi, il termine "eccellente" si riferisce a una percentuale sensibilmente più bassa del 20%, mentre "accettabile" si riferisce a una percentuale ben più alta del 10%. In definitiva, anche un cambio di terminologia gioverebbe ad una migliore accettazione di alcune norme.
- Per quanto la VQR sia stata condotta giustamente area per area, la natura multidisciplinare degli atenei e di molti dipartimenti ha come conseguenza che l'uso della VQR per la distribuzione delle risorse richiede un'integrazione e un confronto di risultati ottenuti in aree diverse. Si ritiene quindi importante per il futuro avere una presentazione dei risultati che normalizzi non soltanto la media ma anche la varianza, in modo da rendere le distribuzioni nelle varie aree meglio confrontabili.
- Una decisione a priori, cruciale per qualsiasi valutazione della ricerca, è: vogliamo valutare la qualità delle eccellenze (come fatto nella VTR 01-03) o la qualità media (come sostanzialmente fatto nella VQR 04-10)? La risposta ha ovviamente ricadute importanti sulle modalità di valutazione. Per esempio, valutare la qualità media implica necessariamente l'esame di un numero molto grande di pubblicazioni, e quindi anche l'uso di tecniche di veloce implementazione, quali quelle bibliometriche, in accompagnamento alla *peer review*. In definitiva, la scelta di come miscelare bibliometria e *peer review*, e di quante pubblicazioni o quali dati richiedere, dipende

fortemente da cosa si vuole valutare. In linea di massima, per avere un quadro completo della situazione scientifica del paese, occorre valutare sia l'eccellenza (da ben definire) che la qualità media. Per farlo tuttavia occorrono tempi, modi e pratiche assai diverse. L'eccellenza, per sua natura, è dei singoli. La risultante dell'eccellenza di molti singoli, può fare (ma non necessariamente fa) l'eccellenza di una struttura. Un'ottima qualità media invece, non assicura affatto la presenza di eccellenze, che possono invece ritrovarsi in strutture con qualità media meno alta. Quindi va sottolineato che la VQR è solo un primo passo verso una *valutazione completa* che l'UMI fortemente auspica: quest'ultima sarà realizzata solo quando disporremo anche di una valutazione della didattica e di dati che mettano in luce come sia realmente distribuita la qualità della ricerca "trainante" del nostro paese.

8) *Possibili valutazioni dei singoli.*

- Una valutazione dei singoli è auspicabile, ma non ha senso se non prevede l'esame competente da parte di una commissione di esperti del curriculum complessivo dell'individuo da valutare. In questo esame si possono considerare anche indicatori quantitativi, come quelli bibliometrici, ma bisogna evitare che la valutazione sia basata in prevalenza su, o peggio si riduca solo all'uso di, questi indicatori.
- Valutazioni con scopi diversi devono essere effettuate con modalità diverse. Limitandosi alla normativa vigente, sono previsti almeno tre tipi principali di valutazione di singoli: quella per le abilitazioni, quella per la (futura e sperabile) attribuzione degli scatti stipendiali, e quella (ancora più lontana e molto auspicabile) per l'attribuzione di incentivi premiali. La prima deve considerare il candidato nel complesso della sua attività scientifica, per giudicarne la maturità; la seconda, rivolta a tutti i docenti, si dovrebbe concentrare sul loro operato degli ultimi anni, e dovrebbe avere come obiettivo quello di escludere dagli scatti solo i casi meno produttivi; la terza infine dovrebbe invece identificare, su domanda, le eccellenze. Ulteriore differenza: la prima è esterna all'ateneo, mentre le ultime due dovrebbero essere interne all'ateneo, anche se è previsto (e auspicabile) l'intervento di valutatori esterni. Non è pensabile utilizzare le stesse modalità nei tre casi: occorre adattare la procedura allo scopo specifico che ci si prefigge, e in tal senso direttive del MIUR e suggerimenti del CUN, dell'ANVUR e della CRUI, sentite le comunità scientifiche, sono auspicabili, onde evitare marchiane differenze e interpretazioni troppo distanti tra i vari atenei.
- Sarà indispensabile avere un modo sistematico, continuativo e uniforme di raccogliere dati sulle pubblicazioni scientifiche dei vari soggetti. Se mai partirà, e l'UMI fortemente auspica che ciò accada, l'ANPrePS (Anagrafe Nazionale Nominativa dei Professori e dei Ricercatori e delle Pubblicazioni Scientifiche) dovrebbe svolgere esattamente questo ruolo: fornire una banca dati nazionale certificata delle pubblicazioni scientifiche di tutti i docenti, in modo da dare in ogni momento una fotografia complessiva e aggiornata della produzione scientifica dei docenti (dei dipartimenti, degli atenei). Va tuttavia sottolineato che spetta alla comunità scientifica nel suo complesso, e non a interventi ad essa esterni, la definizione del concetto cruciale di "pubblicazione scientifica". Dunque a tale riguardo è importante fare riferimento al documento del CUN (cfr. http://www.cun.it/media/123067/proposta_cun_criteri_scientificita.pdf) contenente una proposta sui criteri identificanti il carattere scientifico delle pubblicazioni e degli altri prodotti della ricerca, proposta che fa seguito ad un'ampia e seria consultazione della comunità scientifica (cfr. http://www.cun.it/media/123073/rapporto_esiti_consultazione_pubblica_cun.pdf).
- Come detto, la valutazione periodica dei singoli dovrebbe spettare agli atenei e non all'ANVUR, cui si affiderebbe in caso contrario un compito davvero immane e decisamente troppo gravoso. Scopo dell'ANVUR è infatti la valutazione del sistema universitario nel suo complesso, e non dei singoli. È opportuno invece che siano gli atenei ad assumersi la responsabilità di valutare i propri docenti, in base a criteri generali che vanno stabiliti, e di essere poi valutati per le loro scelte. È opportuno tuttavia che l'ANVUR, in fase di valutazione delle strutture, acquisisca le valutazioni dei singoli effettuate dall'ateneo, per esaminare la qualità e la giustezza degli standard valutativi delle strutture stesse.
- Infine le valutazioni individuali non devono limitarsi a considerare solo gli aspetti della ricerca

scientifico, della didattica e organizzativi, ma anche le attività di terza missione (tra cui, attività di divulgazione, contatti con il mondo della scuola, servizi a vantaggio della comunità scientifica, ecc.).

Questo documento, approvato dalla Commissione Scientifica dell'UMI in data 11/12/2013 è stato redatto con il contributo di Marco Abate, Nicola Bellomo, Paolo Freguglia, Nicola Fusco, Giuseppe Mulone.

