



Appendice A.

Il confronto tra valutazione *peer* e valutazione bibliometrica

A.1. Il campione casuale

Un campione casuale di 1984 articoli su rivista passibili di valutazione bibliometrica è stato estratto dalla popolazione di 21191 articoli, passibili di valutazione bibliometrica, sottomessi alla valutazione nel GEV06. La popolazione è stratificata in base alla distribuzione dei prodotti all'interno dei 4 sub-GEV identificati nell'Area, ossia: Medicina sperimentale, Scienze cliniche, Scienze chirurgiche e Sanità pubblica. La classificazione degli articoli all'interno del sub-GEV si basa sul settore di effettiva valutazione dell'articolo ed è calcolata escludendo i casi di articoli duplicati presentati da diversi autori all'interno di uno stesso strato campionario. Il campione include il 9,5% dei prodotti di Medicina sperimentale, 9,2% di prodotti di Scienze cliniche, 9,6% di prodotti di Scienze chirurgiche e 9,6% di prodotti di Sanità pubblica (Tabella A.1). Il campione è stato estratto ai primi di settembre 2012, prima dell'inizio del processo di revisione *peer*, mediante una procedura casuale con il vincolo di selezionare una proporzione significativa di prodotti in ciascun sub-GEV.

Sub-GEV	Popolazione	Campione	%
Medicina sperimentale	3651	347	9,5
Scienze cliniche	10578	968	9,2
Scienze chirurgiche	5767	554	9,6
Sanità pubblica	1195	115	9,6
Totale	21191	1984	9,4

Tabella A.1: Distribuzione degli articoli su rivista nel campione e nella popolazione

Classe	Popolazione	%	Campione	%
Medicina sperimentale				
E	1672	45,80	148	42,65
B	657	18,00	72	20,75
A	149	4,08	10	2,88
L	506	13,86	47	13,54
IR	667	18,27	70	20,17
Scienze cliniche				
E	4736	44,77	424	43,80
B	1775	16,78	185	19,11
A	455	4,30	39	4,03
L	1607	15,19	154	15,91
IR	2005	18,95	166	17,15
Scienze chirurgiche				
E	1709	29,63	155	27,98

B	843	14,62	78	14,08
A	357	6,19	22	3,97
L	1521	26,37	174	31,41
IR	1337	23,18	125	22,56
Sanità pubblica				
E	430	35,98	47	40,87
B	198	16,57	19	16,52
A	68	5,69	4	3,48
L	259	21,67	24	20,87
IR	240	20,08	21	18,26
Totale				
E	8547	40,33	774	39,01
B	3473	16,39	354	17,84
A	1029	4,86	75	3,78
L	3893	18,37	399	20,11
IR	4249	20,05	382	19,25

Tabella A.2: Distribuzione delle valutazioni bibliometriche nel campione e nella popolazione

La Tabella A.2 riporta la distribuzione nelle classi di valutazione VQR (Eccellente, Buono, Accettabile, Limitato, Incerto (IR)) ottenuta utilizzando la valutazione bibliometrica degli articoli su rivista nelle quattro aree dei sub-GEV, per il campione e per la popolazione. Come si può osservare la distribuzione delle valutazioni bibliometriche (E/ B/ A/ L/ IR) è sufficientemente vicina nella popolazione e nel campione, sia per il totale sia per i sub-GEV così da concludere che il campione estratto è rappresentativo della popolazione di riferimento.

Per ciascun articolo su rivista incluso nel campione casuale sono disponibili le seguenti informazioni:

- Rapporto del primo revisore (P1)
- Rapporto del secondo revisore (P2)
- Rapporto di un eventuale terzo e quarto revisore (P3 e P4)
- Valutazione di sintesi dei giudizi del primo e secondo revisore (P)
- Valutazione bibliometrica (F)

Le variabili P e P1-P4 assumono come valore una delle 4 classi di valutazione E, B, A, L; la valutazione bibliometrica F ha come possibile risultato anche la classe di valutazione “IR”, ossia il suggerimento di procedere con la “*informed peer review*” nel caso di risultati molto diversi tra i due indicatori bibliometrici (Impact Factor e numero citazioni, cfr. i criteri del GEV descritti nell’Appendice B). Le quattro classi, secondo il Bando VQR, sono definite con riferimento ai percentili della distribuzione della qualità degli articoli pubblicati nel mondo.



In particolare, la qualifica di eccellente corrisponde a un articolo che si colloca nel 20% superiore della distribuzione della qualità degli articoli pubblicati nel mondo, quella di buono nel successivo 20%, di accettabile nel successivo 10% e infine quella di limitato nel 50% inferiore. Le variabili P1-P4 sono originariamente misurate su una scala numerica compresa tra 3 a 27, con un punteggio da 1 a 9 assegnato a 3 diversi criteri; tali punteggi sono successivamente utilizzati per determinare per ciascun prodotto sottomesso a valutazione la classe di valutazione *peer* del prodotto, sulla base dei criteri fissati dal GEV¹; le variabili P ed F sono invece rispettivamente espresse in termini delle 4 o 5 classi di valutazione sopra elencate. Sulla base del Bando VQR, alle quattro classi E, B, A, L corrispondono rispettivamente i punteggi 1; 0,8; 0,5; 0.

La classificazione adottata nell'analisi bibliometrica si basa sui criteri descritti nell'APPENDICE B di questo rapporto. Nella revisione dei pari, ai revisori esterni è stato richiesto di valutare ciascun prodotto sulla base della loro percezione soggettiva della qualità del prodotto rispetto alla distribuzione mondiale dei prodotti della ricerca nel settore scientifico a cui il prodotto faceva riferimento. La valutazione dei revisori è stata quindi sintetizzata sulla base di un algoritmo specifico al GEV06, secondo il quale, rispettivamente: i prodotti di classe E erano quelli con un punteggio complessivo compreso tra 23 e 27; i prodotti di classe B avevano un punteggio complessivo compreso tra 18 e 22; i prodotti di classe A un punteggio complessivo compreso tra 15 e 17 e i prodotti di classe L un punteggio compreso tra 3 e 14. Al fine di confrontare i risultati della valutazione bibliometrica e della revisione tra pari, si procede nel seguito a confrontare gli indicatori F e P. Anche altri confronti possono essere tuttavia di importanza significativa: in particolare il confronto tra le valutazioni tra pari P1 e P2 consente di valutare il grado di corrispondenza dei giudizi tra i due revisori².

A.2. Le distribuzioni F e P

Le distribuzioni F e P sopra descritte non sono immediatamente confrontabili, dato che la distribuzione F delle valutazioni bibliometriche comprende una classe IR che non è invece prevista nella valutazione dei pari. E' però possibile ipotizzare che una discordanza di almeno due classi tra la valutazione del primo e secondo revisore segnali un'incertezza della revisione dei pari del tutto analoga a quella che emerge dal confronto tra numero di citazioni e fattore di impatto della sede di pubblicazione nell'analisi bibliometrica: in analogia con la classificazione IR della valutazione bibliometrica, si è creata dunque una corrispondente classificazione "Incerta Peer" (IP) per la valutazione dei pari, al fine di consentire il confronto tra le distribuzioni F e P. La Tabella A.3 mostra la distribuzione in numeri assoluti e percentuali degli indicatori F e P sopra descritti per il totale del campione del GEV06.

¹ L'etichetta "P1", "P2", "P3" e "P4" assegnata ai due revisori è puramente convenzionale e riflette esclusivamente l'ordine di accettazione della proposta di revisione avanzata al potenziale revisore.

² Nel GEV06, è stato necessario procedere anche ad una terza e quarta valutazione di pari in 83 e 4 casi rispettivamente; tali valutazioni non sono state considerate in questa fase di analisi.

Valutazione Bibliometrica (F)	Valutazione Peer (P)					
	E	B	A	L	IP	Totale
E	154	306	74	71	169	774
% rispetto al totale delle valutazioni bibliometriche in classe E	19,90	39,53	9,56	9,17	21,83	100
B	17	123	48	66	100	354
% rispetto al totale delle valutazioni bibliometriche in classe B	4,80	34,75	13,56	18,64	28,25	100
A	3	19	7	29	17	75
% rispetto al totale delle valutazioni bibliometriche in classe A	4,00	25,33	9,33	38,67	22,67	100
L	7	44	37	207	105	400
% rispetto al totale delle valutazioni bibliometriche in classe L	1,75	11,00	9,25	51,75	26,25	100
IR	9	98	39	120	115	381
% rispetto al totale delle valutazioni bibliometriche IR	2,36	25,72	10,24	31,50	30,18	100
Totale	190	590	205	493	506	1984
% rispetto al totale delle valutazioni bibliometriche	9,58	29,74	10,33	24,85	25,50	100

Tabella A.3: Confronto tra F e P – totale del campione

Gli elementi sulla diagonale principale della Tabella A.3 corrispondono ai casi in cui la valutazione dei pari e quella bibliometrica coincidono. Gli elementi al di fuori della diagonale principale corrispondono invece ai casi di non coincidenza tra F e P, o perché la valutazione F è migliore della P (elementi al di sopra della diagonale principale) o viceversa (elementi al di sotto della diagonale). La Tabella A.3 mostra che nella maggior parte dei casi la discordanza tra la valutazione bibliometrica e quella dei pari è dovuta al fatto che la valutazione bibliometrica è più generosa. In particolare, gli articoli classificati come eccellenti sulla base degli indicatori bibliometrici sono 774, contro i soli 190 prodotti eccellenti della valutazione tra pari: solo circa il 20% degli articoli classificati come E secondo la bibliometria ottiene E anche secondo la revisione tra pari, mentre rispettivamente nel 40%, 10% e 9% dei casi i prodotti bibliometricamente eccellenti risultano buoni, adeguati o limitati nella valutazione tra pari. D'altro lato, il numero di articoli che sono classificati in B, A e L dalla valutazione tra pari (590, 205 e 493 articoli rispettivamente) è nettamente più elevato rispetto agli articoli che risultano in B, A ed L secondo la valutazione bibliometrica (354, 75 e 400 articoli rispettivamente). Infine, la numerosità di valutazioni incerte è maggiore nella revisione tra pari (506 articoli) rispetto a quella bibliometrica (381 articoli). Le valutazioni bibliometriche incerte hanno in circa il 28% dei casi una valutazione almeno pari a B nell'analisi *peer*, mentre in circa il 53% dei casi le valutazioni incerte *peer* ricevono un punteggio bibliometrico almeno pari a B.

Complessivamente, l'analisi bibliometrica e la revisione tra pari coincidono nel 31% dei casi. Se si sommano alle valutazioni coincidenti quelle che differiscono di una sola classe, si arriva al 54% del campione. Gli articoli con valutazioni che differiscono per due classi sono 187, il 9% del campione, quelli con massima discordanza (ossia, che differiscono per 3 classi) sono 78 (il 4% del campione). Il restante 33% del campione ha una assegnazione IR o IP con uno dei due metodi, e perviene ad una classe di assegnazione definita secondo l'altro metodo.

La Tabella A.4 mostra la distribuzione degli indicatori P1 e P2. Le valutazioni dei due revisori coincidono nel 38% dei casi, sono diverse per una classe di valutazione nella stessa percentuale dei casi e divergono invece rispettivamente per 2 o 3 classi di valutazione nel 21% e nel 5% dei casi. E' da notare anche che le valutazioni su un giudizio di assegnazione alla classe E sono convergenti in 220 casi, pari a circa il 40% del totale delle valutazioni eccellenti fornite dal primo revisore e dal secondo revisore.

P1	P2				
	E	B	A	L	Totale
E	128	116	40	44	328
% rispetto alle valutazioni del primo revisore in classe E	39,02	35,37	12,20	13,41	100
B	116	291	116	162	685
% rispetto alle valutazioni del primo revisore in classe B	16,93	42,48	16,93	23,65	100
A	40	130	62	121	353
% rispetto alle valutazioni del primo revisore in classe A	11,33	36,83	17,56	34,28	100
L	48	172	134	264	618
% rispetto alle valutazioni del primo revisore in classe L	7,77	27,83	21,68	42,72	100
Totale	332	709	352	591	1984
% rispetto alle valutazioni del primo revisore	16,73	35,74	17,74	29,79	100

Tabella A.4: Confronto tra le valutazioni P1 e P2 – totale del campione

Le Tabelle A.5 e A.6 estendono i risultati delle Tabelle A.3 e A.4 ai singoli sub-GEV. In particolare, dall'analisi dei dati della Tabella A.5 emerge che in tutti i sub-GEV il numero di valutazioni eccellenti è nettamente maggiore secondo la valutazione bibliometrica rispetto alla *peer*. D'altra parte, il numero di valutazioni buone e accettabili è in genere maggiore secondo l'analisi *peer* rispetto a quella bibliometrica.

Medicina sperimentale						
Valutazione Bibliometrica (F)	Valutazione Peer (P)					
	E	B	A	L	IP	Totale
E	37	54	19	13	25	148
% rispetto al totale delle valutazioni bibliometriche in classe E	25,0	36,5	12,8	8,8	16,9	100
B	5	26	16	9	16	72
% rispetto al totale delle valutazioni bibliometriche in classe B	6,9	36,1	22,2	12,5	22,2	100
A	0	4	1	2	3	10
% rispetto al totale delle valutazioni bibliometriche in classe A	0,0	40,0	10,0	20,0	30,0	100
L	1	7	7	19	13	47
% rispetto al totale delle valutazioni bibliometriche in classe L	2,1	14,9	14,9	40,4	27,7	100
IR	1	22	9	16	22	70
% rispetto al totale delle valutazioni bibliometriche IR	1,4	31,4	12,9	22,9	31,4	100
Totale	44	113	52	59	79	347
% rispetto al totale delle valutazioni bibliometriche	12,7	32,6	15,0	17,0	22,8	100
Scienze cliniche						
Valutazione Bibliometrica (F)	Valutazione Peer (P)					
	E	B	A	L	IP	Totale
E	94	169	31	39	91	424
% rispetto al totale delle valutazioni bibliometriche in classe E	22,2	39,9	7,3	9,2	21,5	100
B	9	68	19	39	50	185
% rispetto al totale delle valutazioni bibliometriche in classe B	4,9	36,8	10,3	21,1	27,0	100
A	0	10	5	17	7	39
% rispetto al totale delle valutazioni bibliometriche in classe A	0,0	25,6	12,8	43,6	17,9	100
L	2	16	14	86	37	155
% rispetto al totale delle valutazioni bibliometriche in classe L	1,3	10,3	9,0	55,5	23,9	100
IR	5	38	18	56	48	165
% rispetto al totale delle valutazioni bibliometriche IR	3,0	23,0	10,9	33,9	29,1	100
Totale	110	301	87	237	233	968
% rispetto al totale delle valutazioni bibliometriche	11,4	31,1	9,0	24,5	24,1	100

Scienze chirurgiche						
Valutazione Bibliometrica (F)	Valutazione Peer (P)					
	E	B	A	L	IP	Totale
E	17	60	20	16	42	155
% rispetto al totale delle valutazioni bibliometriche in classe E	11,0	38,7	12,9	10,3	27,1	100
B	3	26	10	12	27	78
% rispetto al totale delle valutazioni bibliometriche in classe B	3,8	33,3	12,8	15,4	34,6	100
A	2	4	1	10	5	22
% rispetto al totale delle valutazioni bibliometriche in classe A	9,1	18,2	4,5	45,5	22,7	100
L	4	16	13	94	47	174
% rispetto al totale delle valutazioni bibliometriche in classe L	2,3	9,2	7,5	54,0	27,0	100
IR	3	33	11	39	39	125
% rispetto al totale delle valutazioni bibliometriche IR	2,4	26,4	8,8	31,2	31,2	100
Totale	29	139	55	171	160	554
% rispetto al totale delle valutazioni bibliometriche	5,2	25,1	9,9	30,9	28,9	100
Sanità pubblica						
Valutazione Bibliometrica (F)	Valutazione Peer (P)					
	E	B	A	L	IP	Totale
E	6	23	4	3	11	47
% rispetto al totale delle valutazioni bibliometriche in classe E	12,8	48,9	8,5	6,4	23,4	100
B	0	3	3	6	7	19
% rispetto al totale delle valutazioni bibliometriche in classe B	0,0	15,8	15,8	31,6	36,8	100
A	1	1	0	0	2	4
% rispetto al totale delle valutazioni bibliometriche in classe A	25,0	25,0	0,0	0,0	50,0	100
L	0	5	3	8	8	24
% rispetto al totale delle valutazioni bibliometriche in classe L	0,0	20,8	12,5	33,3	33,3	100
IR	0	5	1	9	6	21
% rispetto al totale delle valutazioni bibliometriche IR	0,0	23,8	4,8	42,9	28,6	100
Totale	7	37	11	26	34	115
% rispetto al totale delle valutazioni bibliometriche	6,1	32,2	9,6	22,6	29,6	100

Tabella A.5: Confronto tra F e P per sub-GEV

Complessivamente, la tendenza della valutazione bibliometrica a essere più vantaggiosa rispetto a quella peer è comune a tutti i sub-GEV (un test formale di tale ipotesi è presentato nella sezione A.3) I casi di discordanza di due classi o più tra le due valutazioni sono più frequenti in Sanità pubblica (17%) e Medicina sperimentale (14%).

La Tabella A.6 estende i risultati della Tabella A.4 ai singoli sub-GEV. In particolare, dall'analisi emerge che la convergenza nelle valutazioni dei due revisori è relativamente più elevata in Scienze cliniche (39% dei casi), Sanità pubblica e Medicina Sperimentale (37%). Le divergenze per almeno 2 classi di valutazione sono particolarmente frequenti in Sanità pubblica (30%), sub-GEV caratterizzato quindi da un numero più basso di divergenze di una classe. E' da notare anche che le valutazioni su un giudizio di assegnazione alla classe E sono convergenti in almeno la metà dei casi soltanto in Medicina Sperimentale.

Medicina Sperimentale					
P1	P2				
	E	B	A	L	Totale
E	35	16	8	10	69
% rispetto al totale delle valutazioni di classe E del primo revisore	50,72	23,19	11,59	14,49	100
B	25	54	30	23	132
% rispetto al totale delle valutazioni di classe B del primo revisore	18,94	40,91	22,73	17,42	100
A	8	34	13	13	68
% rispetto al totale delle valutazioni di classe A del primo revisore	11,76	50,00	19,12	19,12	100
L	6	24	20	28	78
% rispetto al totale delle valutazioni di classe L del primo revisore	7,69	30,77	25,64	35,90	100
Totale	74	128	71	74	347
% rispetto al totale delle valutazioni del primo revisore	21,33	36,89	20,46	21,33	100
Scienze Cliniche					
P1	P2				
	E	B	A	L	Totale
E	72	65	20	19	176
% rispetto al totale delle valutazioni di classe E del primo revisore	40,91	36,93	11,36	10,80	100
B	59	156	53	76	344
% rispetto al totale delle valutazioni di classe B del primo revisore	17,15	45,35	15,41	22,09	100
A	16	54	26	67	163
% rispetto al totale delle valutazioni di classe A del primo revisore	9,82	33,13	15,95	41,10	100

L	19	83	59	124	285
% rispetto al totale delle valutazioni di classe L del primo revisore	6,67	29,12	20,70	43,51	100
Totale	166	358	158	286	968
% rispetto al totale delle valutazioni del primo revisore	17,15	36,98	16,32	29,55	100
Scienze chirurgiche					
P1	P2				
	E	B	A	L	Totale
E	16	30	11	13	70
% rispetto al totale delle valutazioni di classe E del primo revisore	22,86	42,86	15,71	18,57	100
B	28	60	30	52	170
% rispetto al totale delle valutazioni di classe B del primo revisore	16,47	35,29	17,65	30,59	100
A	13	32	18	38	101
% rispetto al totale delle valutazioni di classe A del primo revisore	12,87	31,68	17,82	37,62	100
L	16	55	42	100	213
% rispetto al totale delle valutazioni di classe L del primo revisore	7,51	25,82	19,72	46,95	100
Totale	73	177	101	203	554
% rispetto al totale delle valutazioni del primo revisore	13,18	31,95	18,23	36,64	100
Sanità pubblica					
P1	P2				
	E	B	A	L	Totale
E	5	5	1	2	13
% rispetto al totale delle valutazioni di classe E del primo revisore	38,46	38,46	7,69	15,38	100
B	4	21	3	11	39
% rispetto al totale delle valutazioni di classe B del primo revisore	10,26	53,85	7,69	28,21	100
A	3	10	5	3	21
% rispetto al totale delle valutazioni di classe A del primo revisore	14,29	47,62	23,81	14,29	100
L	7	10	13	12	42
% rispetto al totale delle valutazioni di classe L del primo revisore	16,67	23,81	30,95	28,57	100
Totale	19	46	22	28	115
% rispetto al totale delle valutazioni del primo revisore	16,52	40,00	19,13	24,35	100

Tabella A.6: Confronto tra le valutazioni P1 e P2 – suddivise per sub-GEV.



A.3. Il confronto tra le distribuzioni di F e P

Il confronto tra la valutazione dei pari e quella bibliometrica si può basare su due criteri fondamentali:

1. Grado di concordanza tra la distribuzione F e la distribuzione P, ossia se F e P tendono ad assegnare lo stesso punteggio ad ogni articolo
2. Grado di differenza sistematica esistente tra F e P misurata mediante la differenza media del punteggio assegnato da F e P sulla base dei pesi attribuiti alle classi della VQR.

Ovviamente, una perfetta concordanza implica anche la non esistenza di differenze sistematiche tra F e P, ma il contrario non è necessariamente vero, e in generale i due criteri misurano due diversi aspetti della differenza esistente tra le due distribuzioni. Si consideri ad esempio una distribuzione con un basso grado di concordanza tra F e P (molti articoli ricevono differenti valutazioni F e P). Anche in tale caso può accadere che, in media, F e P forniscano un punteggio complessivo simile. Questa distribuzione sarebbe caratterizzata da un basso livello di concordanza e da un basso grado di differenza sistematica: adottare uno dei due metodi di valutazione (per esempio quella bibliometrica, F) comporterebbe una frequente differenza di valutazione degli articoli sulla base della bibliometria e della valutazione *peer* (ossia, si avrebbero molti articoli con una buona valutazione in base a F, ma una peggiore valutazione in base a P, o viceversa).

Alternativamente, si consideri un caso di elevata (ma non perfetta) concordanza tra F e P. In questo caso, potrebbe ancora succedere che, per esempio, il numero di articoli con classificazione elevata sia sistematicamente maggiore in F che in P. In questo caso si avrebbe un elevato grado di concordanza, ma anche un alto grado di differenza sistematica tra le due distribuzioni, dato che il punteggio medio attribuito da F differirebbe dal punteggio medio di P. Adottare uno dei due metodi di valutazione può risultare in una sopravvalutazione (o sottovalutazione) in relazione all'altro criterio: ossia, gli articoli riceverebbero un punteggio notevolmente diverso se valutati con F o con P.

Da un punto di vista statistico, il grado di concordanza tra F e P può essere misurato utilizzando la statistica K di Cohen; differenze sistematiche tra F e P possono invece essere misurate guardando alle differenze tra le medie delle distribuzioni e valutandone la significatività con un test *t* di Student.

A.3.1. Il grado di concordanza tra le distribuzioni F e P

La tabella A.7 riporta i valori della statistica K di Cohen, calcolati per l'intero campione e separatamente per ciascun sub-GEV. La statistica K è costruita in modo tale da essere pari a zero quando la concordanza tra due (o più) valutazioni è del tutto casuale, vale a dire nel caso in cui le valutazioni siano indipendenti l'una dall'altra; la statistica assume invece valore pari ad 1 nel caso in cui ci sia perfetta concordanza. E' possibile calcolare il test utilizzando una

matrice standard di pesi lineari (1; 0,67; 0,33; 0) attribuiti ai casi di concordanza, discordanza di una classe e così via, rispettivamente. Le analisi riportate sono riferite a campioni omogenei (*paired sample*), ossia ai prodotti del campione per i quali sono disponibili sia i dati della valutazione *peer* sia quelli relativi alla valutazione bibliometrica, eliminando cioè dal campione i prodotti per i quali la valutazione bibliometrica fornisce come risultato una classificazione IR. In questo caso, nel totale del campione, K è uguale a 0,30, un valore statisticamente diverso da zero agli usuali livelli di significatività. Il grado di concordanza nei 4 sub-GEV presenta valori vicini a quelli registrati per l'intero campione, e comunque sempre statisticamente diversi da zero ad un livello di confidenza dell'1%.

Come accennato sopra, il calcolo di K riportato nella prima riga della tabella usa pesi lineari. E' possibile argomentare che nel nostro caso i pesi appropriati da utilizzare debbano però essere quelli suggeriti dalle regole della VQR. Sulla base di questi pesi, è possibile calcolare la distanza tra le valutazioni utilizzando i punteggi numerici della VQR (1, 0,8, 0,5, 0), associati con le valutazioni qualitative (E, B, A, L). La seconda riga della Tabella A.7 riporta i valori della statistica K calcolati utilizzando i pesi della VQR. In questo caso, il grado di concordanza tra le valutazioni F e P risulta superiore rispetto a quello calcolato utilizzando pesi lineari, sia nel totale del campione che in ciascun sub-GEV.

Test/Area	Totale campione	Medicina Sperimentale	Scienze Cliniche	Scienze chirurgiche	Sanità pubblica
F e P, pesi lineari	0.3030 (19.21)**	0.2407 (6.76)**	0.2883 (13.48)**	0.3368 (10.33)**	0.2023 (3.31)**
F e P, pesi VQR	0.3351 (19.04)**	0.2602 (6.67)**	0.3128 (13.13)	0.3850 (10.59)**	0.2176 (3.13)**
P1 e P2, pesi lineari	0.2466 (13.52)**	0.2320 (5.52)**	0.2791 (10.82)**	0.1821 (5.04)**	0.1474 (2.06)*
P1 e P2, pesi VQR	0.2356 (12.22)**	0.1986 (4.57)**	0.2650 (9.71)**	0.1886 (4.91)**	0.1296 (1.71)*

Nota: La tabella riporta la statistica K e in parentesi il valore z ad essa associato³. La presenza di un asterisco indica la significatività del test al livello dell'5%, la presenza di due asterischi la significatività dell'1%.

Tabella A.7: Statistica K di Cohen sul grado di concordanza

La tabella A.7 riporta anche la statistica K per il grado di concordanza tra i due revisori (P1 e P2), sia per il totale del campione che per le singole aree di ricerca. Nel complesso del campione, il grado di concordanza tra la valutazione bibliometrica (F) e la revisione *peer* (P) è leggermente superiore a quello esistente tra i giudizi formulati tra i due revisori esterni: la

³ Il test z verifica se K è statisticamente pari a zero assumendone una distribuzione Gaussiana, o normale. Si calcola dividendo il valore di K per il suo errore standard. Se il valore di z è superiore al valore-soglia della distribuzione normale corrispondente a una certa probabilità, si conclude che la statistica K è statisticamente diversa da zero, ossia che le due valutazioni non sono indipendenti l'una dall'altra e mostrano quindi un grado statisticamente significativo di concordanza.



statistica K calcolata con pesi lineari è pari a 0,25 (0,24 con i pesi della VQR). Nei singoli sub-GEV il grado di concordanza tra le due valutazioni *peer* si conferma significativo: la statistica K calcolata con pesi lineari assume valori pari rispettivamente a 0,20, 0,26, 0,19, 0,13 rispettivamente, nei sub-GEV di Medicina Sperimentale, Scienze cliniche, Scienze chirurgiche e Sanità pubblica. Il test z ad essa associato conduce sempre a rifiutare l'ipotesi nulla di non concordanza ad un livello di significatività dell'1%, tranne nel sub-GEV di Sanità pubblica, valore significativo al 5%. Risultati analoghi si trovano utilizzando i pesi della VQR al posto dei pesi lineari.

A.4. Il grado di differenza sistematica tra le distribuzioni F e P

La Tabella A.8 riporta il punteggio medio risultante dalle valutazioni F e P. I valori numerici sono ottenuti sommando i pesi assegnati dalla VQR alle quattro classi di merito (rispettivamente, 1; 0.8; 0.5; 0) e dividendo per il numero degli articoli valutati. Si noti ancora una volta come, date le regole della VQR, gli scarti tra F e P non abbiano lo stesso peso: ad esempio, la differenza tra L e A ha un peso di 0,5, mentre la differenza tra E e B ha un peso pari solo a 0,2. I punteggi riportati sono riferiti a campioni omogenei (*paired sample*), ossia ai prodotti del campione per i quali sono disponibili sia i dati della valutazione *peer* sia quelli relativi alla valutazione bibliometrica, eliminando cioè dal campione i prodotti per i quali la valutazione bibliometrica fornisce come risultato una classificazione IR.

La terza colonna mostra che il punteggio medio finale della revisione *peer* (punteggio P) è pari a 0,523: il punteggio medio limitato ai sub-GEV Scienze chirurgiche è inferiore alla media complessiva, mentre è superiore in Medicina sperimentale, Scienze cliniche e Sanità pubblica. Le differenze tra i sub-GEV che emergono dall'analisi dei dati della terza colonna della tabella possono essere attribuite:

- A una migliore qualità degli articoli sottomessi alla valutazione in Medicina Sperimentale e Scienze Cliniche rispetto a quella degli articoli di Scienza Chirurgiche e Sanità Pubblica
- A una maggiore generosità dei revisori dei primi due sub-GEV rispetto a quelli degli ultimi due
- All'intrinseca variabilità statistica nella scelta del campione.

La quarta colonna contiene il punteggio medio ottenuto nella valutazione bibliometrica: tale punteggio è pari a 0,683 per la media complessiva dei lavori sottoposti al GEV06, risultando inferiore alla media generale in Scienze chirurgiche (0,532), uguale alla media in Sanità pubblica (0,683) e superiore ad essa in Medicina sperimentale (0,760) e Scienze cliniche (0,737). L'ordinamento dei sub-GEV in base alla qualità degli articoli presentati varia a seconda che si consideri l'analisi bibliometrica oppure quella *peer*: secondo l'algoritmo bibliometrico, le valutazioni migliori le ricevono gli articoli di Medicina Sperimentale,

seguiti da quelli di Scienze cliniche, scienze chirurgiche e Sanità pubblica. Secondo la valutazione dei pari, invece, le valutazioni migliori sono assegnate sempre a Medicina sperimentale, seguita però da Sanità pubblica, Scienze cliniche e Scienze chirurgiche. Il risultato più interessante dell'analisi è mostrato nella quinta colonna, che presenta la differenza tra valutazione *peer* e bibliometrica, con le colonne 7-8 che riportano il risultato del test *t* per campioni di uguale ampiezza ad esso associato. Nel totale del campione, emerge una differenza sistematica tra la valutazione bibliometrica e la valutazione *peer*: più precisamente, la valutazione media ottenuta con l'analisi bibliometrica è migliore di quella ottenuta con la valutazione *peer*. Il risultato è confermato anche per i dati riferiti ai tutti i 4 sub-GEV.

Sub-GEV	Punteggio P1	Punteggio P2	Punteggio P	Punteggio F	Diff. F-P	# osservazioni	test t	p-value
Medicina sperimentale	0,628	0,632	0,595	0,760	0,165	277	7,079	0,000
Scienze cliniche	0,569	0,571	0,542	0,737	0,195	803	14,179	0,000
Scienze chirurgiche	0,461	0,489	0,436	0,532	0,096	429	4,826	0,000
Sanità pubblica	0,504	0,619	0,558	0,683	0,125	94	2,808	0,006
Totale	0,547	0,562	0,523	0,683	0,160	1603	15,938	0,000

Tabella A.8: Test *t* sulla differenza tra i punteggi bibliometrici e *peer review*

A.5. Conclusioni

Nel totale del campione dei prodotti del GEV06 conferiti per la valutazione, si riscontra una più che adeguata concordanza tra valutazioni effettuate con il metodo della revisione tra pari e con quello bibliometrico. Inoltre, il grado di concordanza tra valutazione finale bibliometrica e *peer* è molto simile al grado di concordanza tra le due valutazioni *peer*. Soltanto nel sub-GEV di Sanità pubblica la concordanza tra due valutazioni *peer* risulta più bassa ai limiti della significatività.

D'altro lato, però, emerge evidenza di differenze sistematiche tra i punteggi corrispondenti alle valutazioni *peer* e bibliometriche. In effetti, è possibile osservare che il numero di prodotti della ricerca classificati come eccellenti (E) con l'algoritmo di valutazione bibliometrica sia superiore a quello dei prodotti "eccellenti" secondo la valutazione tra pari.

Il grado di concordanza tra valutazioni *peer* e valutazioni bibliometriche è elevato in tutti i sub-GEV. Le differenze sistematiche tra i punteggi medi sono statisticamente significative e sempre di segno positivo (ossia, la valutazione bibliometrica è significativamente più favorevole in media rispetto a quella *peer*).



Appendice B.

Documento sui criteri dell'Area 06

Valutazione della Qualità della Ricerca 2004- 2010 (VQR 2004-2010)

Criteri per la valutazione dei prodotti di ricerca

Gruppo di Esperti della Valutazione dell'area 06 Scienze Mediche (GEV06)

Codici European Research Council (ERC): Life Sciences LS1-LS9

1. Delimitazione dell'area GEV06 e Obiettivi della VQR 2004-2010

L'obiettivo finale dell'esercizio di Valutazione della Qualità della Ricerca (VQR) 2004-2010 è quello di fornire un panorama qualitativo della ricerca biomedica in Italia. Nel complesso, i prodotti di ricerca che il GEV06 esaminerà per questo esercizio sono principalmente articoli su riviste scientifiche. Questi articoli saranno classificati prevalentemente utilizzando indicatori bibliometrici integrati allo scopo di valutare la qualità della rivista in cui sono pubblicati e il loro impatto, quantificato in base al numero di citazioni ricevute dall'articolo fino al 31 Dicembre 2011.

2. Organizzazione del GEV06

2.1 Composizione dei Sub-GEV

Il GEV06 è organizzato in 4 sottogruppi (sub-GEV06) così denominati:

1. **Medicina Sperimentale (Experimental Medicine)** che include i macrosettori (Decreto Ministeriale 29 luglio 2011 n. 336) **06/A-Patologia e Diagnostica di Laboratorio** (Settori Scientifico Disciplinari-SSD MED/03-Genetica Medica, MED/04-Patologia Generale, MED/05-Patologia Clinica, MED/46-Scienze Tecniche di Medicina di Laboratorio, MED/02-Storia della Medicina, MED/07-Microbiologia e Microbiologia Clinica, MED/08-Anatomia Patologica).

Coordinatore Prof. Tullio Pozzan.



Componenti: Paolo Gasparini, Alfonso Bellacosa, Riccardo Della Favera, Pier Paolo Pandolfi, Claudio Napoli, Vincent Calvez, Giorgio Palu', Leonardo Antonio Sechi, Giulio Gabbiani, Gaetano Thiene, Federico Cozzolino e Vito Michele Fazio.

2. Scienze cliniche (Clinical Science) che include i macrosettori **06/B–Clinica Medica Generale** (SSD MED/09-Medicina Interna, M-EDF/01-Metodi e Didattiche delle Attività Motorie, M-EDF/02-Metodi e Didattiche delle Attività Sportive), **06/D–Clinica Medica Specialistica** (SSD MED/10-Malattie dell'apparato Respiratorio, MED/11-Malattie dell'Apparato Cardiovascolare, MED/13-Endocrinologia, MED/49-Scienze Tecniche Dietetiche Applicate, MED/14-Nefrologia, MED/15-Malattie del Sangue, MED/16-Reumatologia, MED/06-Oncologia Medica, MED/12-Gastroenterologia, MED/17-Malattie Infettive, MED/35-Malattie Cutanee e Veneree, MED/25-Psichiatria, MED/26-Neurologia), **06/G–Clinica Pediatrica** (SSD MED/38-Pediatria Generale e Specialistica, MED/39-Neuropsichiatria Infantile), **06/I–Clinica Radiologica** (SSD MED/36-Diagnostica per Immagini e Radioterapia, MED/37-Neuroradiologia), **06/L–Clinica Anestesiologica** (SSD MED/41-Anestesiologia), il SSD MED/34-Medicina Fisica e Riabilitativa e il SSD MED/48-Scienze Infermieristiche e Tecniche Neuro-Psichiatriche e Riabilitative.

Coordinatore Prof. Giuseppe Paolisso.

Componenti: Clara Camaschella, Lorenzo Salvatore Malatino, Antonio Pinto, Camillo Ricordi, Francesco Violi, Marina Saetta, Jawahar L. Mehta, Peter Schwartz, Antonio Craxì, Marta Korbonitis, Furio Pacini, Loreto Gesualdo, Carlo Croce, Francesco Lauria, Francesco Lo Coco, Marco Matucci-Cerinic, Francesco Castelli, Thomas Becker, Michele Tansella, Paolo Calabresi, Paolo Maria Rossini, Sandro Sorbi, Luigi Tesio, Ketty Peris, Giuseppe Basso, Andrea Biondi, Luigi Notarangelo, Lieven Lagae, Lorenzo Egildo Derchi, Maria Cristina Messa, Marco Salvatore, Antonio Pesenti, Giovanni Pinto.

3. Scienze Chirurgiche (Surgical Science) che include i macrosettori **06/C–Clinica Chirurgica Generale** (SSD MED/18-Chirurgia Generale), **06/E–Clinica Chirurgica Specialistica** (SSD MED/22-Chirurgia Vascolare, MED/23-Chirurgia Cardiaca, MED/21-Chirurgia Toracica, MED/19-Chirurgia Plastica, MED/20-Chirurgia Pediatrica e Infantile, MED/24-Urologia, MED/27-Neurochirurgia, MED/29-Chirurgia Maxillofacciale), **06/F–Clinica Chirurgica Integrata** (SSD MED/28-Malattie Odontostomatologiche, MED/30-Malattie Apparato Visivo, MED/31-Otorinolaringoiatria, MED/32-Audiologia, MED/33-Malattie Apparato Locomotore), **06/H–Clinica Ginecologica** (SSD MED/40-Ginecologia e Ostetricia), il SSD MED/47-Scienze Infermieristiche Ostetrico-ginecologiche e il SSD MED/50-Scienze Tecniche Mediche Applicate.

Coordinatore Prof. Rocco Domenico Bellantone.

Componenti: Alfonso Barbarisi, Alfred Cuschieri, Peter Friend, Jan Lerut, Paolo Miccoli, Mario Morino, Francesco Moschella, Thomas A. D'Amico, Uberto Bortolotti, Gyorgy Szabò, Michael A.S. Jewett, Giulio Maira, Marco Ferrari, Lorenzo Lo Muzio, Antonella



Polimeni, Stefano Bonini, Leonardo Mastropasqua, Gaetano Paludetti, Robert Ruben, Nicola Maffulli, Domenico Arduini, Felice Petraglia, Giovanni Scambia.

4. Sanità Pubblica (Public Health) che include i macrosettori **06/M–Sanità Pubblica** (SSD MED/42-Igiene Generale e Applicata, MED/01-Statistica Medica, MED/43-Medicina Legale, MED/44-Medicina del Lavoro, MED/45-Scienze Infermieristiche Generali, Cliniche e Pediatriche).

Coordinatore Prof. Antonio Mutti.

Componenti: Giorgio Brandi, Alberto Izzotti, Madea Burkhard, Vincenzo Pascali, Piero Maestrelli, Adriano Decarli.

2.2 Allocazione dei prodotti all'interno del GEV06

L'allocazione dei prodotti ai sub-GEV06 avverrà sulla base del Settore Scientifico Disciplinare (SSD) e della *Subject Category (SC)* del *Journal Citation Reports (JCR)* di *Web of Knowledge (WoK)* indicati dal soggetto valutato così come trasmesso dalle strutture. Il significato del SSD associato al prodotto, che può anche essere diverso dal SSD di appartenenza del soggetto valutato, si riferisce al GEV che con maggior competenza, secondo il soggetto valutato, può valutare il prodotto stesso. Nel caso in cui lo stesso prodotto sia assegnato a più GEV per il suo carattere interdisciplinare, saranno adottati identici criteri di valutazione concordati tra i vari GEV. A tale scopo, i Presidenti dei GEV interessati istituiscono specifici Gruppi di Consenso Inter-Area per concordare il GEV competente.

2.3 Regole di funzionamento del GEV06

Le regole di funzionamento stabilite dal Consiglio direttivo dell'ANVUR sono di sotto richiamate:

- a) il Presidente convoca il GEV06 nel corso della VQR;
- b) La convocazione del GEV06 avviene almeno quindici giorni prima della riunione. La riunione è convocata tramite e-mail dal Presidente, che fissa anche l'ordine del giorno;
- c) Le decisioni all'interno del GEV06 vengono prese a maggioranza semplice dei presenti. Per partecipare alla votazione non è necessario essere fisicamente presenti alle riunioni, purché presenti in modalità telematica.
- d) Alle riunioni del GEV06 partecipa, con funzioni di segretario verbalizzante senza diritto di voto, anche l'assistente al GEV06, Dott.ssa Rosa Margherita Lacatena, assegnata da ANVUR.
- e) Al termine di ciascuna riunione viene redatto un resoconto della seduta in lingua italiana e un estratto succinto contenente le principali decisioni approvate in lingua italiana e in inglese. Il resoconto e il verbale sono inviati ai componenti del GEV06 e approvati tramite e-mail o utilizzando l'ambiente SW predisposto dal CINECA.



3. Mix valutativo

Fermo restando che la responsabilità della valutazione conclusiva è affidata all'intero GEV, il GEV06 utilizzerà per la valutazione la metodologia della *informed peer review*, ovvero un mix di criteri bibliometrici e di revisione *peer* secondo i seguenti criteri:

- a) I prodotti di cui alle tipologie *a* della sezione 2.3 del Bando ANVUR VQR 2004-2010 identificati nelle banche dati *Web of Science (WoS)* di Thomson Reuters verranno valutati utilizzando i criteri bibliometrici descritti nella sezione 3.2. La scelta di *WoS* come base di dati è dovuta al suo preponderante e consolidato utilizzo presso la comunità scientifica internazionale ai fini della valutazione del livello qualitativo della rivista (*Impact Factor*) su cui pubblicare i prodotti della ricerca in campo biomedico.
- b) I prodotti valutati mediante *peer review* (che prevede l'invio ad almeno due revisori esterni oppure, in alcuni casi limitati per i quali vi siano le competenze richieste all'interno, la valutazione diretta da parte di almeno due componenti del GEV06) appartengono a 4 categorie:
 - i. I prodotti di cui alle tipologie *a* della sezione 2.3 del Bando ANVUR VQR 2004-2010 non presenti nella banca dati *WoS*;
 - ii. Articoli (tipologia *a* della sezione 2.3 del Bando ANVUR VQR 2004-2010) che sono indicizzati nella banca dati *WoS* e che richiedono la *peer review* in base all'algoritmo bibliometrico descritto nella sezione 3.2;
 - iii. Articoli (tipologia *a* della sezione 2.3 del Bando ANVUR VQR 2004-2010) che sono indicizzati nella banca dati *WoS* e che saranno valutati utilizzando l'algoritmo bibliometrico e la *peer review* al fine di studiare la correlazione tra i due metodi di valutazione: tali articoli saranno individuati tramite un algoritmo di campionamento casuale stratificato studiato da un Gruppo di lavoro dell'ANVUR.
 - iv. I prodotti delle tipologie *b, c, d, e* elencate nella sezione 2.3 del Bando ANVUR VQR 2004-2010;
- c) La scelta dei prodotti da inviare a *peer review* terrà conto delle raccomandazioni formulate dalle strutture nell'apposito campo dell'interfaccia CINECA e della tipologia del prodotto (prodotti provenienti da attività di ricerca in aree emergenti a livello internazionale o in aree di forte specializzazione o a carattere interdisciplinare, in ragione della minore rappresentazione di tali aree negli indicatori bibliometrici).
- d) La valutazione dei prodotti attribuiti dalla struttura ai soggetti valutati non terrà conto della posizione del soggetto nella lista degli autori ma terrà conto unicamente della qualità del prodotto. Tuttavia, data l'importanza della posizione dell'autore nelle consuetudini pubblicistiche dell'area 06 Scienze Mediche, il rapporto conclusivo dell'area 06 (Rapporto Finale GEV06) riporterà, per i prodotti valutati nella classe di merito Eccellente e per ciascuna struttura, la percentuale di prodotti nei quali il



soggetto valutato, cui è attribuito il prodotto, risulta come primo autore, ultimo autore, o autore corrispondente. Si suggerisce pertanto alle strutture, ove possibile, di attribuire i prodotti a soggetti che, come indicato dalla loro posizione nella lista degli autori, abbiano contribuito in maniera preminente al lavoro.

3.1 Peer review

Nel caso di utilizzo della *peer review* i prodotti saranno inviati a due revisori esterni oppure valutati, sussistendo le competenze e le condizioni di assenza di conflitti, da parte di due componenti del GEV06 utilizzando la medesima scheda di revisione.

La selezione dei revisori esterni, italiani e stranieri, attese le sue rilevanti finalità di pubblico interesse, si informa al principio di leale cooperazione istituzionale ed è retta da criteri di correttezza, obiettività e imparzialità. Ciascun componente del GEV06 è tenuto a garantire l'anonimato degli esperti cui è affidata la valutazione dei prodotti. Le informazioni acquisite ai fini della selezione dei revisori possono essere utilizzate esclusivamente ai fini della valutazione della qualità della ricerca.

La scelta dei revisori esterni avverrà evitando conflitti di interesse tra i revisori stessi e gli autori e/o la/le struttura/e di affiliazione. Inoltre, verrà garantita l'indipendenza dei revisori ponendo attenzione alla sede di affiliazione, alle collaborazioni scientifiche recenti o in atto e alla nazionalità. Per minimizzare i problemi legati ai potenziali conflitti di interesse, si privilegeranno i revisori operanti in Università ed istituzioni straniere. L'individuazione di ciascuno dei due revisori *peer* esterni verrà fatta indipendentemente da due componenti distinti del sub-GEV06 di riferimento, ove possibile.

3.1.1 L'individuazione dei revisori *peer*

Il GEV06 intende coinvolgere revisori esterni con un curriculum scientifico di alto profilo secondo standard internazionali testimoniato, in particolare negli ultimi anni, da un elevato numero di pubblicazioni nelle sedi di riferimento della comunità scientifica internazionale del settore, un significativo numero di citazioni e la necessaria esperienza di valutazione nella specifica area. Il GEV06 preparerà un elenco di revisori esterni, stabilendo standard minimi di qualità scientifica, di impatto sulla comunità scientifica internazionale e di esperienza nella valutazione.

Grande attenzione verrà posta al mantenimento dell'anonimato dei revisori, sia nella fase di predisposizione dell'elenco dei revisori, che nella fase operativa di valutazione.

Per quanto attiene alla prima, il Presidente del GEV06 consulterà la lista di revisori della propria area resa disponibile dal CINECA, e chiederà ai componenti del GEV06, tramite i quattro coordinatori dei sub-GEV06, di suggerire un ampio numero di esperti che soddisfano i criteri sopra indicati.

Il Presidente del GEV06 raccoglierà le indicazioni corredate di informazioni fornite sulla base di una scheda condivisa, e, anche con l'ausilio dei coordinatori dei sub-GEV06, provvederà a modificare la lista CINECA con integrazioni e/o cancellazioni.



Il processo di integrazione della lista continuerà per tutta la durata della valutazione, sulla base delle necessità che dovessero emergere a valle della trasmissione dei prodotti da parte delle strutture.

3.1.2 Assegnazione dei livelli di merito VQR sulla base delle valutazioni peer

Al fine di garantire l'omogeneità dei giudizi dei revisori e la loro adeguata motivazione, il GEV06 predisporrà una apposita scheda revisore costituita da una serie di domande a risposta multipla e da un campo libero con numero limitato di parole. Il GEV06 trasformerà le indicazioni contenute nella scheda revisore in una delle quattro classi finali di merito.

Nel caso di valutazioni non convergenti dei due revisori *peer* esterni, dei due componenti del GEV06 operanti come revisori interni o tra *peer review* e analisi bibliometrica, il sub-GEV06 creerà al suo interno un Gruppo di Consenso costituito da almeno tre componenti (o di numero maggiore ma sempre dispari) con il compito di proporre al GEV06 il punteggio finale del prodotto oggetto del giudizio discordante fra revisori e/o fra revisori e classificazione bibliometrica mediante la metodologia del *consensus report*. Il Gruppo di Consenso del sub-GEV06 formulerà la sua proposta a maggioranza semplice dei componenti. Il Gruppo di Consenso potrà avvalersi anche del giudizio di un terzo esperto. In ogni caso la responsabilità della valutazione conclusiva è dell'intero GEV06, non del sub-GEV06 né del componente che appartiene ad uno specifico SSD. A tale scopo, il GEV06 terrà conto delle valutazioni dei componenti del sub-GEV06, delle revisioni *peer*, della classificazione su base bibliometrica e delle proposte del Gruppo di Consenso. Nel caso di pubblicazioni prive di indicatori bibliometrici quali per esempio prodotti sotto forma editoriale di Libro, capitolo di Libro, o pubblicazione multimediale, si terrà conto anche delle caratteristiche dell'Edizione: per esempio, collana editoriale nella quale una monografia è stata pubblicata, l'esistenza di un comitato editoriale, adozione di procedure trasparenti di revisione per decidere sulla pubblicazione, diffusione e prestigio a livello nazionale e internazionale dei prodotti dell'Editore, recensioni dell'opera pubblicate su riviste internazionali, e ogni altro elemento atto a fornire indicazioni utili circa la qualità e l'impatto scientifici dell'opera.

3.2 Analisi bibliometrica

3.2.1 Basi di dati

Il GEV06 utilizzerà *Web of Science* di Thomson Reuters come base di dati. La scelta di *WoS* come base di dati è dovuta al suo preponderante e consolidato utilizzo presso la comunità scientifica internazionale ai fini della valutazione del livello qualitativo della rivista (*Impact Factor*) su cui pubblicare i prodotti della ricerca in campo biomedico.



3.2.2 Finestra temporale delle citazioni

Nel calcolo dell'indicatore bibliometrico il GEV06 utilizzerà le citazioni fino al 31 dicembre 2011.

3.2.3 Auto-citazioni

L'opportunità di includere o escludere le autocitazioni nella valutazione bibliometrica è tuttora oggetto di dibattito nella comunità scientifica. Nella VQR 2004-2010, per motivi di ordine tecnico legati soprattutto all'uso di *WoS* che non consente la diretta distinzione tra etero-citazioni e auto-citazioni e a causa di problemi legati alla disambiguazione dei nomi degli autori a partire dai dati grezzi in possesso di CINECA, non verranno escluse le autocitazioni.

3.2.4 Gli indicatori bibliometrici

La valutazione utilizzerà, per tutti gli articoli pubblicati su riviste indicizzate nella base di dati *WoS*, un algoritmo che tiene conto, in misura diversa a seconda della data di pubblicazione dell'articolo, sia del numero di citazioni che dell'indicatore bibliometrico della rivista ospitante (*Impact Factor*).

Tale scelta è dettata dalle seguenti considerazioni:

- a. Il puro indicatore citazionale assume valori che possono essere anche molto piccoli in dipendenza della disciplina e dell'età dell'articolo, rendendo difficile una discriminazione effettiva tra classi di merito; queste considerazioni sono il risultato di una significativa sperimentazione effettuata utilizzando le basi di dati acquisite per l'esercizio VQR 2004-2010;
- b. L'uso del solo indicatore citazionale costituisce un obiettivo facile per possibili future manipolazioni, e potrebbe indurre comportamenti fuorvianti quali l'estensiva autocitazione e la citazione mutua all'interno di un gruppo ristretto, non giustificate da considerazioni di natura tecnica e scientifica;
- c. Si ritiene importante comunicare ai giovani ricercatori il messaggio che esiste un livello qualitativo anche profondamente diverso tra le varie riviste e che ci si debba cimentare con severe procedure di *peer review* e avere, quale obiettivo, quello di pubblicare nelle riviste più selettive del proprio settore.

Di seguito sono indicati i criteri seguiti e l'algoritmo utilizzato per la classificazione degli articoli nelle quattro classi di merito della VQR 2004-2010:

1. Dato l'articolo e la rivista che lo ha pubblicato, si identifica la corrispondente *SC* in *JCR* di *WoK* considerando congiuntamente sia prodotti inseriti nella *Science Edition*, sia prodotti inseriti nella *Social Science Edition* di *JCR*;
2. Se la rivista appartiene a più di una *SC*, si utilizza, ai fini dell'individuazione univoca della *SC*, l'indicazione del soggetto valutato che ha proposto l'articolo o, se necessario, l'eventuale modifica da parte del GEV;

3. Per garantire una maggiore rappresentatività e robustezza degli indicatori bibliometrici, il GEV06 su proposta dei sub-GEV ha provveduto alla fusione di alcune *Subject Categories (SCs)* di *JCR*, in particolare per:
 - Categorie che includono meno di quaranta riviste;
 - Categorie che hanno una marcata affinità scientifica e culturale e condividono molte riviste con altre *SCs*.

L'elenco delle *SCs* e degli accorpamenti è allegato al presente documento (**Allegato 1**).

4. La *Subject Category* “Multidisciplinary Science” che include riviste caratterizzate da una pluralità di argomenti scientifici, quali *Nature*, *Science*, *PNAS* e alcune *SCs* che coprono argomenti inter-disciplinari (vedi **Allegato 1**) saranno valutate usando questa procedura:

Gli articoli pubblicati su una rivista di tali *SCs* saranno riassegnati ad un'altra *SC* sulla base dei riferimenti bibliografici prevalenti contenuti nell'articolo. In particolare, per ognuna delle riviste citate si individuerà una (o più) *SC* di appartenenza, e si sceglierà la *SC* finale mediante la metodologia del *consensus report* sopra riportata che ne individui quella prevalente. Nell'assegnazione alla nuova *SC*, l'articolo manterrà l'*Impact Factor (IF)* assegnato alla rivista stessa e il numero di citazioni ricevute;

5. Si calcola la funzione di distribuzione cumulativa empirica dell'*IF* delle riviste appartenenti a ciascuna delle *SCs* (singole od accorpate come da **Allegato 1**) individuate per l'anno di pubblicazione dell'articolo da valutare;
6. Si divide la funzione di distribuzione cumulativa in quattro classi, caratterizzate dai valori di frequenza 0.2 (classe 1), 0.2 (classe 2), 0.1 (classe 3), 0.5 (classe 4), ottenendo i 3 valori soglia di *IF* che separano le suddette classi;
7. Si calcola la funzione di distribuzione cumulativa empirica del numero di citazioni di tutti gli articoli (dalla data di pubblicazione al 31 dicembre 2011) pubblicati dalle riviste appartenenti alla *SC* individuata per l'anno di pubblicazione dell'articolo da valutare;
8. Si divide la funzione di distribuzione cumulativa del numero di citazioni in quattro classi, caratterizzate dai valori di frequenza 0.2 (classe 1), 0.2 (classe 2), 0.1 (classe 3), 0.5 (classe 4), ottenendo i 3 valori soglia di citazioni che separano le suddette classi;
9. Con i valori ottenuti delle soglie di *IF* e di citazioni per ogni anno di VQR, si costruisce una matrice 4x4 di riferimento che contiene sulle colonne le classi della distribuzione dell'*IF* e sulle righe le classi della distribuzione del numero di citazioni secondo il modello generale mostrato nella Figura 1; tali matrici seguiranno specificatamente lo schema della Matrice I (Figura 2) per gli anni dal 2004 al 2008, mentre seguiranno quello della Matrice II (Figura 3) per il 2009 e 2010;
10. Ottenuti il valore di *IF* della rivista ed il numero di citazioni dell'articolo da valutare, esso viene associato ad una delle sedici coppie di classi presenti nella matrice di riferimento dell'anno di pubblicazione (Matrice I o II);

11. L'attribuzione della classe finale di merito avviene secondo il seguente algoritmo: la lettera "A" si riferisce alla classe finale "Eccellente", la "B" a "Buono", la "C" a "Accettabile", e la "D" a "Limitato":
- Quando le coordinate dell'articolo lo posizionano in una dei quattro elementi della diagonale principale, e quindi le due indicazioni basate su *IF* e su citazioni coincidono, la classe finale è la stessa (vedi Figura 1).

		Indicatore bibliometrico			
		1	2	3	4
n. di citazioni	1	A			
	2		B		
	3			C	
	4				D

Figura 1. Matrice generale di corrispondenza tra classi iniziali di *IF* e citazioni e classe finale VQR

- Quando le coordinate danno indicazioni discordanti (elemento della matrice non appartenente alla diagonale principale), occorre distinguere due casi a seconda della data di pubblicazione dell'articolo.
- Se l'articolo è stato pubblicato nel quinquennio 2004-2008, potendo quindi considerare sufficientemente consolidata la distribuzione di citazioni ad esso associata, si utilizza la matrice riportata in Figura 2 per l'attribuzione della classe finale di merito, nella quale si privilegia l'indicazione fornita dalla classificazione citazionale (vedi Figura 2).

2004-2008

Indicatore bibliometrico

		1	2	3	4
n. di citazioni	1	A	A	A	IR
	2	B	B	B	IR
	3	IR	C	C	C
	4	IR	D	D	D

Figura 2. Matrice I. Matrice di corrispondenza tra classi iniziali di *IF* e citazioni e classe finale VQR per articoli pubblicati nel quinquennio 2004-2008 (IR= *informed peer review*)

- d. Se l'articolo è stato pubblicato nel biennio 2009-2010, e quindi la sua "storia" citazionale è, al contrario, ritenuta non sufficientemente consolidata, si utilizza la matrice seguente per l'attribuzione della classe finale di merito (vedi Figura 3), nella quale si privilegia l'indicazione fornita dalla classificazione basata sull'*IF*.

2009-2010

Indicatore bibliometrico

		1	2	3	4
n. di citazioni	1	A	IR	IR	IR
	2	A	B	C	D
	3	A	B	C	D
	4	IR	IR	IR	D

Figura 3. Matrice II. Matrice di corrispondenza tra classi iniziali di *IF* e citazioni e classe finale VQR per articoli pubblicati nel biennio 2009-2010 (IR= *Informed peer review*)

e. Gli elementi delle matrici di **Figura 2** e **Figura 3** etichettati “IR” si riferiscono ai casi nei quali la valutazione non verrà fatta in maniera automatica, ma mediante la *informed peer review*. In tali casi, il GEV06 valuterà direttamente l’articolo, o, in mancanza al suo interno delle competenze necessarie, lo affiderà alla *peer review* esterna.

12. L’algoritmo distinguerà, calcolando distribuzioni cumulative empiriche separate per le citazioni, gli articoli “scientifici” (*original paper*) da quelli di rassegna (*review*), che ricevono notoriamente un maggior numero di citazioni.

3.3 Gestione dei conflitti

Nel caso di disponibilità di valutazione *peer* e bibliometrica per lo stesso prodotto, eventuali conflitti di attribuzione verranno risolti dal GEV06 su proposta motivata del sub-GEV, sentito il Gruppo di Consenso sopra richiamato.

3.4 Assenza di indicatori di citazione

Tutti i prodotti non contenuti nelle basi di dati citazionali *WoS* e privi di indicatori bibliometrici saranno sottoposti a valutazione diretta da parte di due componenti del GEV06, o, in assenza delle competenze necessarie al suo interno, a *peer review* da parte di revisori esterni selezionati dal GEV06.

4. Altri prodotti

Per quanto concerne i prodotti non appartenenti alla tipologia *a* della sezione 2.3 del Bando ANVUR VQR 2004-2010 il GEV06 si atterrà ai seguenti criteri:

- a) Gli Atti di congresso (*Proceedings*) pubblicati su riviste con codice ISSN o ISBN (tipicamente come Supplementi alle riviste) saranno valutabili soltanto con livello di classificazione VQR non superiore al livello di merito D (Limitato).
- b) Per quanto attiene libri o capitoli su libri, verranno valutati solo prodotti in cui il soggetto valutato rientri come Autore, e non prodotti in cui il soggetto valutato rientri solo come *Editor*- Curatore o come Traduttore.
- c) Prodotti didattici rivolti a studenti di Corsi di Laurea o di Specializzazione non verranno valutati. Prodotti editi in proprio non verranno valutati.
- d) Prodotti in forma di “*Author’s reply*”, “*Commentary*”, “*Erratum*”, “*Corrispondenza*”, “*Obituary*” o recensione di libri e articoli non saranno valutati.
- e) Le pubblicazioni che hanno per oggetto i risultati di “*trial clinici*” o di studi multicentrici, quali ad esempio studi genetici, saranno presi in considerazione nella valutazione solo nel caso in cui il soggetto valutato sia autore del manoscritto mentre la sola partecipazione come sperimentatore attestata dalla presenza nella lista degli sperimentatori/collaboratori o nei ringraziamenti in fondo all’articolo comporta l’esclusione del prodotto dalla valutazione.



- f) I brevetti saranno sottoposti a *peer review* da parte di esperti esterni, anche stranieri. I livelli di merito VQR A (Eccellente) e B (Buono) potranno essere assegnati esclusivamente a brevetti internazionali.

5. Risoluzione dei conflitti di interesse

I membri dei GEV si asterranno dal valutare o dall'assegnare ad altri componenti dei GEV o a esperti esterni:

- a. prodotti di cui siano co-autori;
- b. prodotti di cui siano co-autori loro parenti, congiunti o affini fino al 4° grado;
- c. prodotti presentati da strutture (Enti di ricerca vigilati dal MIUR, Università e Dipartimenti Universitari, soggetti pubblici e privati che svolgono attività di ricerca) presso cui i componenti stessi lavorino tuttora, o abbiano lavorato o svolto incarichi o collaborazioni ufficiali negli ultimi sette anni (a partire dal 1/1/2004);
- d. prodotti presentati da enti di ricerca vigilati dal MIUR e da altri soggetti pubblici e privati sottoposti volontariamente alla VQR presso cui i componenti stessi abbiano o abbiano avuto un rapporto di lavoro o con le quali abbiano svolto incarichi o collaborazioni ufficiali, inclusa l'affiliazione a enti di ricerca, negli anni a partire dal 1/1/2004.

Nei casi di cui al punto d) precedente, esiste conflitto di interesse:

- i. nel caso in cui la struttura abbia una permanente strutturazione interna di tipo territoriale o disciplinare (es. sezione locale di ente di ricerca, istituto, dipartimento), limitatamente ai prodotti presentati dalla stessa articolazione;
- ii. nel caso in cui la struttura non abbia una permanente strutturazione interna di tipo territoriale o disciplinare (es. sezione locale di ente di ricerca, istituto, dipartimento), in riferimento a tutti i prodotti presentati nei limiti in cui ciò sia possibile senza precludere la possibilità di valutare il prodotto;
- iii. nel caso in cui la strutturazione interna abbia luogo a più livelli gerarchici (es. più istituti riuniti sotto un dipartimento) il conflitto di interesse sorge al livello più basso (es. membri GEV affiliati ad istituti diversi di uno stesso dipartimento, sono in conflitto di interesse soltanto rispetto a prodotti presentati da autori appartenenti allo stesso istituto).

Nei casi di conflitti di interesse, il Presidente del GEV06 incaricherà delle procedure di valutazione un altro componente del GEV06 per il quale non vi siano conflitti di interesse.

Nel caso di conflitti di interesse che coinvolgano il Presidente del GEV06, l'assegnazione dei prodotti relativi sarà fatta dal coordinatore della VQR o da persona da lui incaricata.

ALLEGATO 1

L'elenco delle *Subject Categories* sottoposte ad accorpamento:

Accorpamenti	Subject Category
1	ALLERGY
1	IMMUNOLOGY
2	ANDROLOGY
2	ENDOCRINOLOGY & METABOLISM
3	ANESTHESIOLOGY
3	CRITICAL CARE MEDICINE
3	EMERGENCY MEDICINE
4	CARDIAC & CARDIOVASCULAR SYSTEMS
4	PERIPHERAL VASCULAR DISEASE
5	INFECTIOUS DISEASES
5	TROPICAL MEDICINE
6	HEMATOLOGY
6	RHEUMATOLOGY
7	GERIATRICS & GERONTOLOGY
7	GERONTOLOGY (Social Science)
7	MEDICINE, GENERAL & INTERNAL
7	PRIMARY HEALTH CARE
8	CRIMINOLOGY & PENOLOGY (Social Science)
8	MEDICINE, LEGAL
9	REHABILITATION (Social Science)
9	REHABILITATION
10	PUBLIC, ENVIRONMENTAL & OCCUPATIONAL HEALTH (Social Science)
10	PUBLIC, ENVIRONMENTAL & OCCUPATIONAL HEALTH
10	ERGONOMICS (Social Science)
11	NURSING
11	NURSING (Social Science)
12	PSYCHIATRY (Social Science)
12	PSYCHIATRY

13	MICROBIOLOGY
13	VIROLOGY
13	PARASITOLOGY
13	MYCOLOGY
14	HEALTH CARE SCIENCES & SERVICES
14	HEALTH POLICY & SERVICES (Social Science)
15	NEUROSCIENCES
15	NEUROIMAGING
16	SUBSTANCE ABUSE
16	PHARMACOLOGY & PHARMACY
17	CELL & TISSUE ENGINEERING
17	MATERIALS SCIENCE, BIOMATERIALS
17	MEDICAL LABORATORY TECHNOLOGY
17	MEDICINE, RESEARCH & EXPERIMENTAL
18	DEVELOPMENTAL BIOLOGY
18	ANATOMY & MORPHOLOGY
18	MICROSCOPY
19	BIOCHEMISTRY & MOLECULAR BIOLOGY
19	BIOPHYSICS
19	POLYMER SCIENCE
20	BIOCHEMICAL RESEARCH METHODS
20	CHEMISTRY, ANALYTICAL
20	MATHEMATICAL & COMPUTATIONAL BIOLOGY
20	MEDICAL INFORMATICS

Subject Categories che coprono argomenti inter-disciplinari:

	Subject Category
1	TRANSPLANTATION
2	REPRODUCTIVE BIOLOGY
3	BIOLOGY
4	CHEMISTRY, MEDICINAL
5	INTEGRATIVE & COMPLEMENTARY MEDICINE
6	MULTIDISCIPLINARY SCIENCES
7	EDUCATION, SCIENTIFIC DISCIPLINES
8	FOOD SCIENCE & TECHNOLOGY



Appendice C.

Schede di valutazione per la peer review

Evaluation form for products (GEV06)

Q1. Relevance: Score the importance of the work, namely the value and impact of the article in advancing the field in terms of scientific knowledge, technical capability, and/or improvement of clinical practice.

Impact	Score	Additional Guidance on Strengths/Weaknesses
High	9	Exceptionally strong with essentially no weaknesses
	8	Extremely strong with negligible weaknesses
	7	Very strong with only some minor weaknesses
Medium	6	Strong but with numerous minor weaknesses
	5	Strong but with at least one moderate weakness
	4	Some strengths but also some moderate weaknesses
Low	3	Some strengths but with at least one major weakness
	2	A few strengths and a few major weaknesses
	1	Very few strengths and numerous major weaknesses
<p>Minor Weakness: An easily addressable weakness that does not substantially lessen impact</p> <p>Moderate Weakness: A weakness that lessens impact</p> <p>Major Weakness: A weakness that severely limits impact</p>		



Q2. Originality / Innovation: Provide a score on the level of originality and innovation of the article. Did the article challenge existing paradigms or clinical practice? Did the article develop novel concepts, approaches/methodologies, tools, or technologies for this area of Medicine?

Impact	Score	Additional Guidance on Strengths/Weaknesses
High	9	Exceptionally strong with essentially no weaknesses
	8	Extremely strong with negligible weaknesses
	7	Very strong with only some minor weaknesses
Medium	6	Strong but with numerous minor weaknesses
	5	Strong but with at least one moderate weakness
	4	Some strengths but also some moderate weaknesses
Low	3	Some strengths but with at least one major weakness
	2	A few strengths and a few major weaknesses
	1	Very few strengths and numerous major weaknesses
<p>Minor Weakness: An easily addressable weakness that does not substantially lessen impact</p> <p>Moderate Weakness: A weakness that lessens impact</p> <p>Major Weakness: A weakness that severely limits impact</p>		



Q3. Internationalization/Impact: Score the degree of international standing of the work in terms of impact, competitiveness and visibility within the international community.

Impact	Score	Additional Guidance on Strengths/Weaknesses
High	9	Exceptionally strong with essentially no weaknesses
	8	Extremely strong with negligible weaknesses
	7	Very strong with only some minor weaknesses
Medium	6	Strong but with numerous minor weaknesses
	5	Strong but with at least one moderate weakness
	4	Some strengths but also some moderate weaknesses
Low	3	Some strengths but with at least one major weakness
	2	A few strengths and a few major weaknesses
	1	Very few strengths and numerous major weaknesses
<p>Minor Weakness: An easily addressable weakness that does not substantially lessen impact</p> <p>Moderate Weakness: A weakness that lessens impact</p> <p>Major Weakness: A weakness that severely limits impact</p>		



Optional:

If deemed useful include a final comment on this product (1-2 sentences, 200-300 characters):

Evaluation form for patents (GEV06)

Q1. Transfer: Score the transfer capability of the invention, namely the commercial and licensing potential

Impact	Score	Additional Guidance on Strengths/Weaknesses
High	9	Exceptionally strong with essentially no weaknesses
	8	Extremely strong with negligible weaknesses
	7	Very strong with only some minor weaknesses
Medium	6	Strong but with numerous minor weaknesses
	5	Strong but with at least one moderate weakness
	4	Some strengths but also some moderate weaknesses
Low	3	Some strengths but with at least one major weakness
	2	A few strengths and a few major weaknesses
	1	Very few strengths and numerous major weaknesses
<p>Minor Weakness: An easily addressable weakness that does not substantially lessen impact</p> <p>Moderate Weakness: A weakness that lessens impact</p> <p>Major Weakness: A weakness that severely limits impact</p>		

Q2. Technology development: Provide a score on the technology development of the invention, considering:

- *time and investment to bring the invention to the market;*
- *whether the invention requires a simple or complex technology to be implemented*

Impact	Score	Additional Guidance on Strengths/Weaknesses
High	9	Exceptionally strong with essentially no weaknesses
	8	Extremely strong with negligible weaknesses
	7	Very strong with only some minor weaknesses
Medium	6	Strong but with numerous minor weaknesses
	5	Strong but with at least one moderate weakness
	4	Some strengths but also some moderate weaknesses
Low	3	Some strengths but with at least one major weakness
	2	A few strengths and a few major weaknesses
	1	Very few strengths and numerous major weaknesses
<p>Minor Weakness: An easily addressable weakness that does not substantially lessen impact</p> <p>Moderate Weakness: A weakness that lessens impact</p> <p>Major Weakness: A weakness that severely limits impact</p>		

Q3. Socio-economic Impact: Score the socio-economic impact of the patent. Is the invention providing an improvement in the prevention and/or the cure of a disease, and/or the quality of patient's life? What is the potential market size of the invention and its ability to capture market share?

Impact	Score	Additional Guidance on Strengths/Weaknesses
High	9	Exceptionally strong with essentially no weaknesses
	8	Extremely strong with negligible weaknesses
	7	Very strong with only some minor weaknesses
Medium	6	Strong but with numerous minor weaknesses
	5	Strong but with at least one moderate weakness
	4	Some strengths but also some moderate weaknesses
Low	3	Some strengths but with at least one major weakness
	2	A few strengths and a few major weaknesses
	1	Very few strengths and numerous major weaknesses
<p>Minor Weakness: An easily addressable weakness that does not substantially lessen impact</p> <p>Moderate Weakness: A weakness that lessens impact</p> <p>Major Weakness: A weakness that severely limits impact</p>		



Optional:

If deemed useful include a final comment on this product (1-2 sentences, 200-300 characters):



Appendice D. Lista delle tabelle con le didascalie

Tabella 1.1: *Settori scientifico disciplinari (SSD) dell'Area 06 (Scienze mediche).*

Tabella 1.2: *Composizione del Gruppo di Esperti della Valutazione - Area 06.* Per ogni componente è indicata l'affiliazione e il ruolo nel Gruppo di Esperti della Valutazione dell'Area 06 (GEV06).

Tabella 1.3: *Organizzazione degli esperti in sub-GEV, corrispondenti SSD e distribuzione dei prodotti della ricerca.* Il GEV06 è organizzato in 4 sub-GEV. Per ogni sub-GEV vengono indicati i corrispondenti SSD e i prodotti gestiti da ogni componente.

Tabella 1.4: *Elenco delle riunioni del GEV 06.*

Tabella 1.5: *Elenco delle strutture valutate dal GEV06.* Nella tabella sono indicate le strutture valutate dal GEV06: 63 Università (U); 3 enti di ricerca (E); 7 consorzi (C).

Tabella 1.6: *Distribuzione dei prodotti di ricerca conferiti per tipologia e SSD.* Per SSD si intende il settore scientifico disciplinare di afferenza del soggetto valutato. La tipologia "Altro" comprende tutte le tipologie di prodotto non menzionate precedentemente. Tali tipologie sono state accorpate a causa del numero ridotto di prodotti che includono. I settori scientifico disciplinari non sono definiti per i prodotti presentati da enti e consorzi (SSD non definito).

Tabella 1.7: *Distribuzione dei prodotti di ricerca per lingua di pubblicazione e SSD.* Per SSD si intende il settore scientifico disciplinare di afferenza del soggetto valutato. La categoria "Altra lingua" contiene le percentuali di prodotti di ricerca pubblicati in lingue diverse da italiano e inglese. La categoria "Lingua non disponibile" contiene le percentuali dei prodotti di ricerca per i quali la lingua di pubblicazione non è stata specificata. I settori scientifico disciplinari non sono definiti per i prodotti presentati da enti e consorzi (SSD non definito).

Tabella 1.8: *Distribuzione dei lavori conferiti per tipologia, lingua di pubblicazione e sub-GEV.* Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV. La tipologia "Altro" comprende tutte le tipologie di prodotto non menzionate precedentemente. Tali tipologie sono state accorpate a causa del numero ridotto di prodotti che includono. La categoria "Altra lingua" contiene le percentuali di prodotti di ricerca pubblicati in lingue diverse da italiano e inglese. La categoria "Lingua non disponibile" contiene le percentuali dei prodotti di ricerca per i quali la lingua di pubblicazione non è stata specificata. I sub-GEV non sono definiti per i prodotti presentati da enti e consorzi (sub-GEV non definito).

Tabella 1.9: Distribuzione dei lavori conferiti per anno di pubblicazione, tipologia e SSD. Per SSD si intende il settore scientifico disciplinare di afferenza del soggetto valutato. La tipologia “Altro” comprende tutte le tipologie di prodotto non menzionate precedentemente. Tali tipologie sono state accorpate a causa del numero ridotto di prodotti che includono. I settori scientifico disciplinari non sono definiti per i prodotti presentati da enti e consorzi (SSD non definito).

Tabella 1.10: Prodotti attesi e conferiti per l'Area 06. Nella tabella è indicato il numero dei prodotti attesi per l'Area 06, il numero dei prodotti sottomessi da ricercatori dell'Area 06 (al GEV06 e ad altri GEV), il numero dei prodotti conferiti al GEV06 da ricercatori di altre aree.

Tabella 1.11: Prodotti conferiti all'Area per tipologia. I prodotti considerati sono quelli conferiti da ricercatori afferenti all'area e comprendono i duplicati.

Tabella 2.1: Distribuzione di revisori e revisioni per nazionalità e SSD. Nella Tabella è indicato per ogni SSD: il numero di revisori Italiani e non Italiani; il numero delle revisioni effettuate da revisori Italiani e non Italiani. Nella tabella non sono comprese le revisioni effettuate dai componenti GEV. E' da notare che uno stesso revisore può essere stato coinvolto nella revisione dei prodotti di più SSD tra di loro affini.

Tabella 2.2: Numero di revisioni effettuate, inevase o rifiutate dai revisori Italiani e Non Italiani. Nella Tabella è indicato il numero di revisioni effettuate, inevase e rifiutate con varie motivazioni dai revisori Italiani e non Italiani. Nella tabella non sono comprese le revisioni effettuate dai componenti GEV.

Tabella 2.3: Distribuzione del numero di revisioni per revisore. Nella Tabella è indicata la distribuzione del numero di revisioni per revisore e la percentuale relativa. Nella tabella non sono comprese le revisioni effettuate dai componenti GEV.

Tabella 2.4: Numero di prodotti conferiti e attesi, per SSD. Per SSD si intende il settore scientifico disciplinare di afferenza del soggetto valutato. Il numero di prodotti attesi è stato calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. Il numero di prodotti conferiti contiene il numero di prodotti effettivamente sottomessi. I settori scientifico disciplinari non sono definiti per i prodotti presentati da enti e consorzi (SSD non definito).

Tabella 2.5: Numero di soggetti valutati e relativo numero di prodotti attesi. SV è l'acronimo di Soggetti Valutati. Il numero di prodotti attesi è stato calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. I settori scientifico disciplinari non sono definiti per i prodotti presentati da enti e consorzi (SSD non definito).



Tabella 2.6: Numero e percentuale di prodotti di ricerca conferiti da soggetti afferenti all'area, per area di valutazione. Per GEV si intende il GEV che ha valutato i prodotti di ricerca.

Tabella 2.7: Percentuale di prodotti di ricerca valutati attraverso la peer review, la bibliometria e con entrambi i metodi di valutazione. Per prodotti valutati in bibliometria si intendono prodotti che hanno una valutazione bibliometrica secondo i criteri del GEV valutante (anche non conclusiva, cioè IR). Per prodotti con doppia valutazione si intendono i prodotti che sono stati sottoposti ad entrambi i metodi di valutazione (peer review e bibliometria). Le percentuali della tabella non tengono conto dei prodotti penalizzati. I settori scientifico disciplinari non sono definiti per i prodotti presentati da enti e consorzi (SSD non definito).

Tabella 2.8: Punteggi ottenuti e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-) nell'area e per sub-GEV. Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV. Per “somma punteggi (v)” si intende la valutazione complessiva del sub-GEV ottenuta sommando i punteggi dei prodotti associati a soggetti valutati dell'area afferenti al sub-GEV. La categoria “prodotti penalizzati” contiene i prodotti ritenuti non valutabili e i casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non sottomessi), i prodotti identici presentati più volte dalla stessa struttura, i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. I sub-GEV non sono definiti per i prodotti presentati da enti e consorzi (sub-GEV non definito).

Tabella 2.9: Punteggi ottenuti e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-) per SSD. Per “somma punteggi (v)” si intende la valutazione complessiva del SSD ottenuta sommando i punteggi dei prodotti afferenti al SSD. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non sottomessi), i prodotti identici presentati più volte dalla stessa struttura, i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. I settori scientifico disciplinari non sono definiti per i prodotti presentati da enti e consorzi (SSD non definito).



Tabella 2.10: Punteggi ottenuti e distribuzione dei prodotti conferiti non penalizzati nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-) per sub-GEV e anno di pubblicazione. Punteggi ottenuti e distribuzione dei prodotti conferiti non penalizzati nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-) per sub-GEV e anno di pubblicazione. Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV. Per “somma punteggi (v)” si intende la valutazione complessiva del sub-GEV per l'anno considerato, ottenuta sommando i punteggi dei prodotti afferenti al sub-GEV.

Tabella 2.11: Punteggi ottenuti e distribuzione dei prodotti conferiti non penalizzati nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-) per modalità di valutazione (indicizzato WoS o non indicizzato WoS). La tabella non considera i prodotti penalizzati. Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti valutati secondo la modalità indicata ottenuta sommando i punteggi dei prodotti valutati.

Tabella 2.12: Punteggi ottenuti e distribuzione dei prodotti conferiti non penalizzati nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-) per tipologia di pubblicazione. Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti alla tipologia indicata, ottenuta sommando i punteggi dei singoli prodotti. La tipologia “Altro” comprende tutte le tipologie di prodotto non menzionate precedentemente. Tali tipologie sono state accorpate a causa del numero ridotto di prodotti che includono.

Tabella 3.1.a: Graduatorie degli atenei piccoli (# prodotti attesi: 10-299) ordinate per voto medio ($I=v/n$) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-). Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti alla struttura ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non inviati), i prodotti identici presentati più volte dalla stessa struttura e i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R indica il voto medio normalizzato. La tabella non include le strutture con meno di 10 prodotti attesi. Gli atenei evidenziati in giallo non hanno un corso di laurea in Medicina e Chirurgia.

Tabella 3.1.b: Graduatorie degli atenei medi (# prodotti attesi: 300-699) ordinate per voto medio ($I=v/n$) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-). Per “somma punteggi (v)” si intende la



valutazione complessiva dei prodotti appartenenti alla struttura ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non inviati), i prodotti identici presentati più volte dalla stessa struttura e i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R indica il voto medio normalizzato.

Tabella 3.1.c: *Graduatorie degli atenei grandi (# prodotti attesi: ≥ 700) ordinate per voto medio ($I=v/n$) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-).* Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti alla struttura ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non inviati), i prodotti identici presentati più volte dalla stessa struttura e i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R indica il voto medio normalizzato.

Tabella 3.2: *Graduatorie degli enti di ricerca ordinate per voto medio ($I=v/n$) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-).* Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti alla struttura ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non inviati), i prodotti identici presentati più volte dalla stessa struttura e i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R indica il voto medio normalizzato. La tabella non include le strutture con meno di 19 prodotti attesi.

Tabella 3.3: *Graduatorie dei consorzi ordinate per voto medio ($I=v/n$) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-).* Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti alla struttura ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a



inviare alla VQR. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non inviati), i prodotti identici presentati più volte dalla stessa struttura e i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università e Consorzio). R indica il voto medio normalizzato. La tabella non include le strutture con meno di 10 prodotti attesi.

Tabella 3.4: Graduatorie di strutture per sub-GEV ordinate per voto medio ($I=v/n$) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-). Per sub-GEV si intende un sottogruppo di SSD interni all’area con affinità disciplinari costituiti secondo i criteri del GEV. Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti alla struttura ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non sottomessi), i prodotti identici presentati più volte dalla stessa struttura e i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). La tabella non include le università e i consorzi con meno di 10 prodotti attesi e gli enti con meno di 19 prodotti attesi. I sub-GEV non sono definiti per i prodotti presentati da enti e consorzi (sub-GEV non definito).

Tabella 3.5: Graduatorie di strutture per SSD ordinate per voto medio ($I=v/n$) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-). Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti alla struttura ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non sottomessi), i prodotti identici presentati più volte dalla stessa struttura e i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). U: Università; E: ente di ricerca; C: Consorzio. I settori scientifico disciplinari non sono definiti per i prodotti presentati da enti e consorzi (SSD non definito).

Tabella 3.6.a: Indicatore IRAS1 per gli atenei piccoli (# prodotti attesi: 10-299). R indica il voto medio normalizzato. $(n/N) \times 100$ rappresenta la percentuale dei prodotti attesi della struttura rispetto ai prodotti attesi dell’area (N: numero prodotti attesi dell’Area; n: numero prodotti attesi della struttura). IRAS1 è il primo indicatore di qualità della ricerca definito nel Bando VQR come rapporto tra il punteggio complessivo



raggiunto da una struttura in una data area e la valutazione complessiva dell'area stessa. La tabella non include le strutture con meno di 10 prodotti attesi. Gli atenei evidenziati in giallo non hanno un corso di laurea in Medicina e Chirurgia.

Tabella 3.6.b: Indicatore IRAS1 per gli atenei medi (# prodotti attesi: 300-699). R indica il voto medio normalizzato. $(n/N) \times 100$ rappresenta la percentuale dei prodotti attesi della struttura rispetto ai prodotti attesi dell'area (N: numero prodotti attesi dell'Area; n: numero prodotti attesi della struttura). IRAS1 è il primo indicatore di qualità della ricerca definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da una struttura in una data area e la valutazione complessiva dell'area stessa.

Tabella 3.6.c: Indicatore IRAS1 per gli atenei grandi (# prodotti attesi: ≥ 700). R indica il voto medio normalizzato. $(n/N) \times 100$ rappresenta la percentuale dei prodotti attesi della struttura rispetto ai prodotti attesi dell'area (N: numero prodotti attesi dell'Area; n: numero prodotti attesi della struttura). IRAS1 è il primo indicatore di qualità della ricerca definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da una struttura in una data area e la valutazione complessiva dell'area stessa.

Tabella 3.7: Indicatore IRAS1 per gli enti di ricerca. R indica il voto medio normalizzato. $(n/N) \times 100$ rappresenta la percentuale dei prodotti attesi della struttura rispetto ai prodotti attesi dell'area (N: numero prodotti attesi dell'Area; n: numero prodotti attesi della struttura). IRAS1 è il primo indicatore di qualità della ricerca definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da una struttura in una data area e la valutazione complessiva dell'area stessa. La tabella non include le strutture con meno di 19 prodotti attesi.

Tabella 3.8: Indicatore IRAS1 per i consorzi. R indica il voto medio normalizzato. $(n/N) \times 100$ rappresenta la percentuale dei prodotti attesi della struttura rispetto ai prodotti attesi dell'area (N: numero prodotti attesi dell'Area; n: numero prodotti attesi della struttura). IRAS1 è il primo indicatore di qualità della ricerca definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da una struttura in una data area e la valutazione complessiva dell'area stessa. La tabella non include le strutture con meno di 10 prodotti attesi.

Tabella 4.1: Graduatorie dei dipartimenti delle Università (post L. 240) ordinati per voto medio ($I=v/n$) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-). Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti alla struttura ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. La categoria “prodotti penalizzati” contiene i



prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non sottomessi), i prodotti identici presentati più volte dalla stessa struttura e i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R rappresenta il voto medio normalizzato. $(n/N) \times 100$ rappresenta la percentuale dei prodotti attesi dal dipartimento rispetto ai prodotti attesi dell'area (N : numero prodotti attesi dell'Area; n : numero prodotti attesi del dipartimento). IRD1 è il primo indicatore di qualità della ricerca applicato ai dipartimenti definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da un dipartimento in una data area e la valutazione complessiva dell'area stessa. La tabella non include i dipartimenti con meno di 10 prodotti attesi.

Tabella 4.2: Graduatorie dei dipartimenti delle Università (post L.240) ordinati per voto medio ($I=v/n$) all'interno delle strutture (elencate in ordine alfabetico) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-). Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti alla struttura ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a sottomettere. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non sottomessi), i prodotti identici presentati più volte dalla stessa struttura e i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R rappresenta il voto medio normalizzato. $(n/N) \times 100$ rappresenta la percentuale dei prodotti attesi dal dipartimento rispetto ai prodotti attesi dell'area (N : numero prodotti attesi dell'Area; n : numero prodotti attesi del dipartimento). IRD1 è il primo indicatore di qualità della ricerca applicato ai dipartimenti definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da un dipartimento in una data area e la valutazione complessiva dell'area stessa. La tabella non include i dipartimenti con meno di 10 prodotti attesi.

Tabella 4.3: Graduatorie delle sottostrutture degli enti di ricerca ordinate per voto medio ($I=v/n$) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-). Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti alla struttura ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non sottomessi), i prodotti identici presentati più volte dalla stessa struttura e i prodotti identici presentati più volte dallo



stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R rappresenta il voto medio normalizzato. $(n/N) \times 100$ rappresenta la percentuale dei prodotti attesi dalla sottostruttura rispetto ai prodotti attesi dell'area (N : numero prodotti attesi dell'Area; n : numero prodotti attesi del dipartimento). IRD1 è il primo indicatore di qualità della ricerca applicato ai dipartimenti definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da un dipartimento in una data area e la valutazione complessiva dell'area stessa. La tabella non include le sottostrutture con meno di 19 prodotti attesi.

Tabella 4.4: Graduatorie dei dipartimenti delle Università (ante L.240) ordinati per voto medio ($I=v/n$) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-). Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti alla struttura ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non sottomessi), i prodotti identici presentati più volte dalla stessa struttura e i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R rappresenta il voto medio normalizzato. $(n/N) \times 100$ rappresenta la percentuale dei prodotti attesi dal dipartimento rispetto ai prodotti attesi dell'area (N : numero prodotti attesi dell'Area; n : numero prodotti attesi del dipartimento). IRD1 è il primo indicatore di qualità della ricerca applicato ai dipartimenti definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da un dipartimento in una data area e la valutazione complessiva dell'Area stessa. La tabella non include i dipartimenti con meno di 10 prodotti attesi.

Tabella 4.5: Graduatorie dei dipartimenti delle Università (ante L.240) ordinati per voto medio ($I=v/n$) all'interno delle strutture (elencate in ordine alfabetico) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-). Graduatorie dei dipartimenti ante L.240 ordinati per voto medio ($I=v/n$) all'interno delle strutture (elencate in ordine alfabetico) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-). Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti alla struttura ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a sottomettere inviare alla VQR. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non sottomessi), i prodotti identici presentati più volte dalla stessa struttura e i prodotti



identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R rappresenta il voto medio normalizzato secondo la formula (4). $(n/N) \times 100$ rappresenta la percentuale dei prodotti attesi dal dipartimento rispetto ai prodotti attesi dell'area (N: numero prodotti attesi dell'Area; n: numero prodotti attesi del dipartimento). IRD1 è il primo indicatore di qualità della ricerca applicato ai dipartimenti definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da un dipartimento in una data area e la valutazione complessiva dell'area stessa. La tabella non include i dipartimenti con meno di 10 prodotti attesi.

Tabella 4.6: Graduatorie dei dipartimenti post L.240 per sub-GEV ordinati per voto medio ($I=v/n$) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-). Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV. Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti al dipartimento ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non sottomessi), i prodotti identici presentati più volte dalla stessa struttura e i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). U: Università; E: ente di ricerca; C: Consorzio. I sub-GEV non sono definiti per i prodotti presentati da enti e consorzi (sub-GEV non definito). La tabella non include i dipartimenti e i consorzi con meno di 10 prodotti attesi e gli enti con meno di 19 prodotti attesi. I sub-GEV non sono definiti per i prodotti presentati da enti e consorzi.

Tabella 4.7: Graduatorie dei dipartimenti post L.240 per SSD ordinati per voto medio ($I=v/n$) e distribuzione dei prodotti nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A, Limitato -L-). Per “somma punteggi (v)” si intende la valutazione complessiva dei prodotti appartenenti al dipartimento ottenuta sommando i punteggi dei singoli prodotti. Per # prodotti attesi si intende il numero di prodotti attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a inviare alla VQR. La categoria “prodotti penalizzati” contiene i prodotti non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non sottomessi), i prodotti identici presentati più volte dalla stessa struttura e i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). U: Università; E: ente di ricerca; C: Consorzio. La tabella non include i dipartimenti e i consorzi con meno di 10 prodotti attesi e gli enti con meno di 19 prodotti attesi.



I settori scientifico disciplinari non sono definiti per i prodotti presentati da enti e consorzi (SSD non definito).

Tabella 4.8: *Graduatoria delle strutture per voto medio (VM) dei soggetti valutati (SV) e distribuzione in classi di voto medio.* Per VM si intende il voto medio dei ricercatori della struttura calcolato come rapporto tra la somma del punteggio medio dei prodotti presentati da ciascun ricercatore e il numero di soggetti valutati della struttura. Le 8 classi di voto medio contengono rispettivamente il numero dei soggetti valutati con $VM < 0$, $VM = 0$, $0 < VM \leq 0.2$, $0.2 < VM \leq 0.4$, $0.4 < VM \leq 0.6$, $0.6 < VM \leq 0.8$, $0.8 < VM \leq 0.99$, $VM = 1$. Inoltre la tabella riporta il numero di soggetti valutati i cui prodotti hanno tutti classe finale E (eccellente) assunti prima e dopo il 2004. U: Università; E: ente di ricerca; C: Consorzio. La tabella non include le strutture con meno di 4 soggetti valutati e con meno di 4 soggetti valutati assunti prima o dopo il 2006 con tutti i prodotti E.

Tabella 4.9: *Graduatoria dei dipartimenti ex L.240 per voto medio (VM) dei soggetti valutati (SV) e distribuzione in classi di voto medio.* Per VM si intende il voto medio dei ricercatori del dipartimento calcolato come rapporto tra la somma del punteggio medio dei prodotti presentati da ciascun ricercatore e il numero di soggetti valutati del dipartimento. Le 8 classi di voto medio contengono rispettivamente il numero dei soggetti valutati con $VM < 0$, $VM = 0$, $0 < VM \leq 0.2$, $0.2 < VM \leq 0.4$, $0.4 < VM \leq 0.6$, $0.6 < VM \leq 0.8$, $0.8 < VM \leq 0.99$, $VM = 1$. Inoltre la tabella riporta il numero di soggetti valutati i cui prodotti hanno tutti classe finale E (eccellente) assunti prima e dopo il 2006. Sulla base del Bando, e per i soli ricercatori delle Università e ricercatori e tecnologi degli enti di ricerca, il numero di prodotti attesi subiva delle riduzioni in funzione dell'anno della prese di servizio nei rispettivi ruoli. Per tutte le altre categorie, non erano previste riduzioni se non legate a determinate categorie di congedo. Nel costruire la tabella, si è identificato l'anno della presa di servizio con il numero di prodotti attesi sulla base del Bando, senza tenere conte di eventuali congedi. U: Università; E: ente di ricerca; C: Consorzio. La tabella non include i dipartimenti con meno di 4 soggetti valutati e con meno di 4 soggetti valutati assunti prima o dopo il 2006 con tutti i prodotti E.

Tabella 5.1: *Punteggio medio dei soggetti valutati per sub-GEV e anno di assunzione in ruolo.* Sulla base del Bando, e per i soli ricercatori delle Università e ricercatori e tecnologi degli enti di ricerca, il numero di prodotti attesi subiva delle riduzioni in funzione dell'anno della prese di servizio nei rispettivi ruoli. Per tutte le altre categorie, non erano previste riduzioni se non legate a determinate categorie di congedo. Nel costruire la tabella, si è identificato l'anno della presa di servizio con il numero di prodotti attesi sulla base del Bando, senza tenere conte di eventuali congedi. Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV. I periodi di presa servizio sono quelli riportati nella Tabella 3 del Bando VQR che stabilisce il numero di prodotti attesi. Il punteggio medio è calcolato come rapporto del punteggio complessivo e il numero dei prodotti attesi. Il numero di prodotti



attesi è stato calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di prodotti che da bando questi erano tenuti a sottomettere. I sub-GEV non sono definiti per i prodotti presentati da enti e consorzi (sub-GEV non definito).

Tabella 5.2: Numero soggetti valutati attivi, non attivi, parzialmente attivi e parzialmente inattivi per struttura. La colonna “# SV” indica il numero di soggetti valutati (SV) della struttura. Per soggetti valutati attivi (SV Attivi) si intendono i soggetti valutati che hanno conferito un numero di prodotti uguale al numero di prodotti attesi. Per soggetti valutati non attivi (SV NA) si intendono i soggetti valutati che non hanno presentato alcun prodotto. Per soggetti valutati parzialmente attivi (SV PA) si intendono i soggetti valutati che non hanno presentato tutti i prodotti attesi ma un numero superiore alla metà. Per soggetti valutati parzialmente inattivi (SV PI) si intendono i soggetti valutati che hanno presentato al più la metà dei prodotti attesi. La colonna “% (NA+PA+PI)” rappresenta la quota di soggetti valutati non pienamente attivi ottenuta considerando i non attivi (NA), i parzialmente attivi (PA) e i parzialmente inattivi (PI). U: Università; E: ente di ricerca; C: Consorzio. La tabella non include i dipartimenti con meno di 4 soggetti valutati.