



## La valutazione dell'efficacia del progetto PQM in italiano

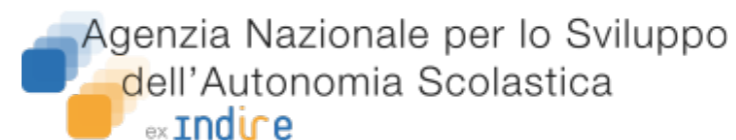
*Antonella Costanzo, Ph.D. - INVALSI*



Unione Europea  
P.O.N. - "Competenze per lo Sviluppo" (FSE)  
D.G. Occupazione, Affari Sociali e pari Opportunità



Ministero dell'Istruzione, dell'Università e della Ricerca  
Dipartimento per la Programmazione  
D.G. per gli Affari Internazionali - Ufficio IV  
Programmazione e gestione dei fondi strutturali europei  
e nazionali per lo sviluppo e la coesione sociale



*Progetto realizzato con il contributo del Fondo Sociale Europeo – Programma Operativo Nazionale "Competenze per lo Sviluppo"  
Azione A.2 "Definizione di Strumenti e metodologie per l'autovalutazione/valutazione del servizio scolastico"*

- PQM e apprendimenti in italiano
- Valutazione dell'“effetto PQM”: questioni preliminari
- Il disegno di valutazione
- I dati PQM e la comprensione della lettura
- Stima dell' impatto di PQM sulle performance in italiano: evidenze
- Riflessioni conclusive

- **Il Piano:** miglioramento degli apprendimenti in italiano. In particolare sviluppo e potenziamento delle competenze in lettura e di grammatica attraverso una didattica innovativa.
- **I destinatari:** *studenti di classi* afferenti a *docenti* assegnati al progetto in *scuole* secondarie di primo grado appartenenti alle 4 Regioni PON per il triennio 2010-13.
- **Struttura:** interventi integrati pluriennali di recupero/potenziamento apprendimenti; condivisione di saperi e di buone pratiche con formazione tra pari per insegnanti; produzione di materiali.

# Valutazione dell'“effetto PQM”: questioni preliminari

## Effetto su cosa?

Performances in italiano misurate in termini di punteggi risultanti da prove standardizzate .

- Unità di analisi: *classi* dello stesso plesso scolastico e della stessa sezione appaiate tra 2 annualità, 2011/12 (pre-intervento) e 2012/13 (post-intervento), distinte in classi partecipanti e classi non partecipanti.

## Effetto di cosa?

Complesso di interventi e misure finalizzate allo sviluppo e all'attuazione di una didattica innovativa per il miglioramento della qualità di insegnamento e delle competenze in italiano.

### La domanda di ricerca:

**L'attuazione di PQM ha migliorato significativamente le performances in italiano per i partecipanti rispetto ai non partecipanti?**

Disegno *quasi-sperimentale* (trattamento non *randomizzato*).

Possibili criticità:

- fenomeno di *selection bias* (Heckman, 1979): possibili differenze di partenza tra le classi trattate e di controllo non dipendenti dal trattamento;
- minacce alla validità interna.

Si utilizza un *approccio controfattuale*: costruzione di un opportuno gruppo di controllo (*post-hoc*) mediante metodi di *matching* (Angrist & Pischke, 2008) sfruttando la ricchezza dei dati SNV.

## Strategia di analisi

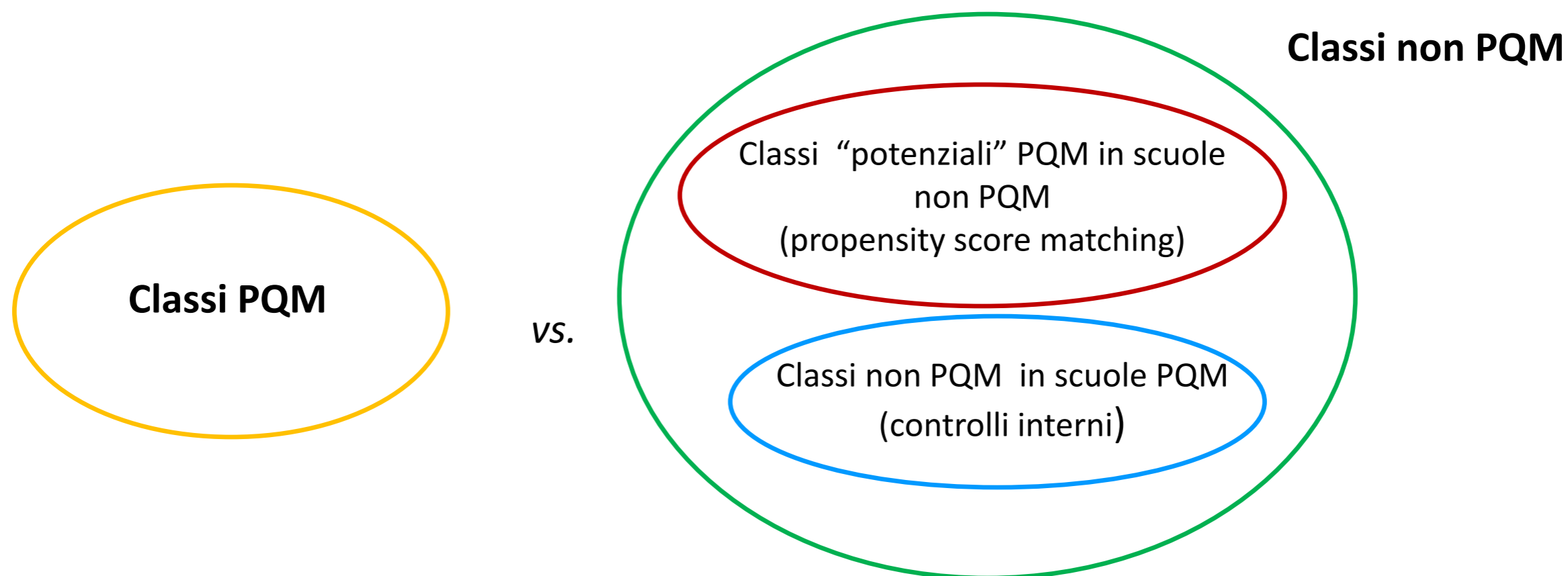
Modello Lineare Generalizzato a Misure Ripetute per la valutazione dell'efficacia del programma sugli apprendimenti in italiano (Field, 2012).

- ✓ Si tiene conto della struttura di dipendenza dei dati (classi *panel* tra le coorti).

- **Idea:** selezionare e abbinare le unità di controllo più *simili* alle unità trattate rispetto al set di variabili osservabili prima del trattamento (Rosembaum & Rubin, 1983).
- **Set di variabili pre-trattamento:** % femmine, % studenti immigrati, % ripetenti, voto medio in italiano, ESCS medio di classe, dimensione classe.
- **Strategia identificazione gruppo di controllo:** *propensity score matching*
  - ✓ Le classi di controllo sono abbinate 1 a 1 alle classi trattate considerando i valori più vicini in termini di *propensity score* (procedura *Nearest Neighbor Matching*).
- **Il *matching*** delle unità avviene stratificando per provincia per preservare la dimensione “territorialità”.

# Il disegno di valutazione dell'efficacia

2 coorti contigue, classi III per a.s. 2011/12 (pre-intervento) e 2012/13 (post intervento) per un totale di 610 classi distinte in classi trattate e di controllo.



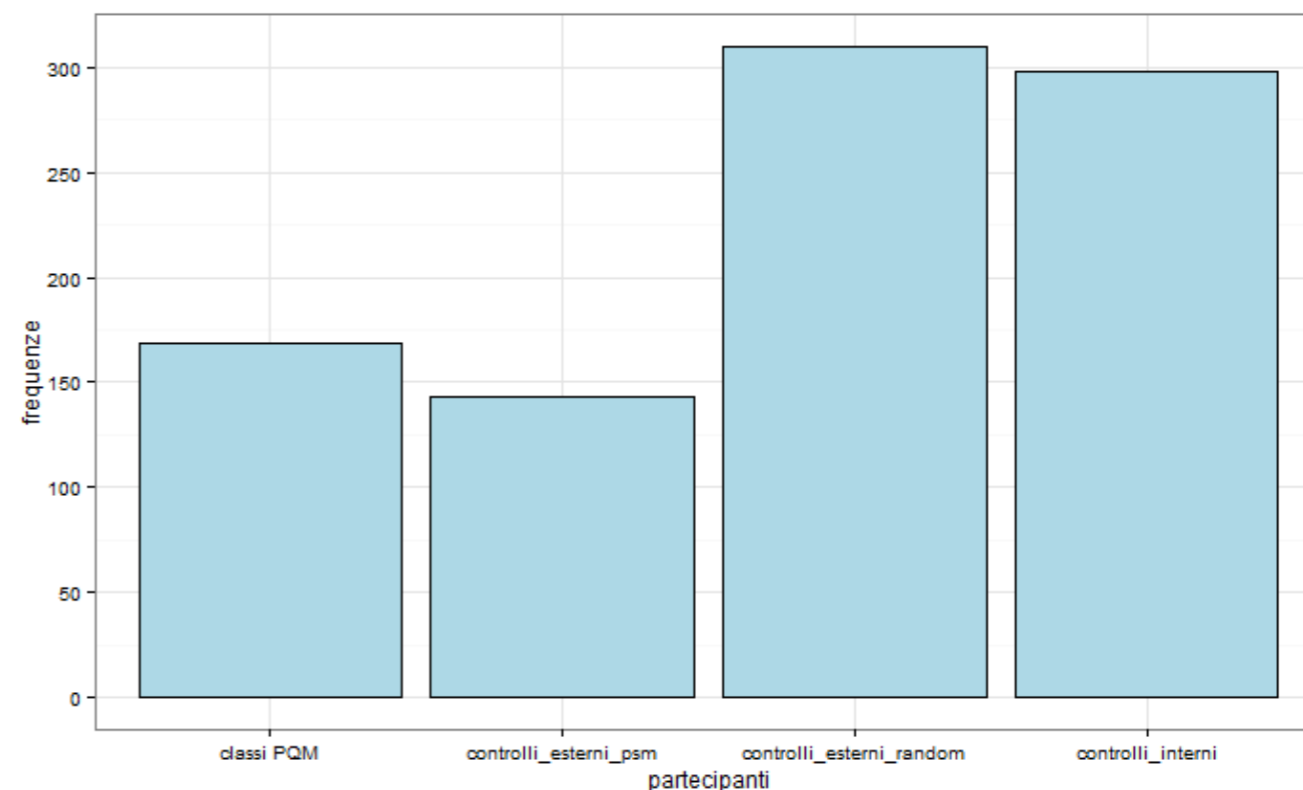
Confronto tra classi PQM e *pool* di classi non partecipanti nell'ambito del quale vengono individuati opportuni gruppi di controllo:

- **“potenziali” PQM:** classi afferenti a scuole non PQM abbinate alle classi trattate mediante procedura di *propensity score matching*;
- **controlli interni:** classi non PQM afferenti a scuole selezionate per partecipare al progetto PQM.

Nell'ambito del disegno di valutazione PQM per le misure degli apprendimenti si sfrutta la ricchezza dei dati SNV (III media, Prova Nazionale per entrambe le coorti).

## I Partecipanti

Gruppo	Numerosità	Frequenza %
Classi PQM	169	27.70
Controlli interni	298	47.68
Controlli esterni (psm)	143	22.88
Totale disegno	610	100
<b>Classi non PQM</b>	310	



**Nota operativa:** nella costruzione del disegno si considerano in via ulteriore tra le classi non PQM, un set di classi selezionate casualmente dal *pool* delle classi disponibili e con simili caratteristiche



### Motivi:

- ✓ Migliorare la validità interna dello studio
- ✓ Maggiore robustezza delle stime sull'impatto di PQM
- ✓ Una più completa caratterizzazione degli effetti di PQM sugli apprendimenti



## Principali caratteristiche e composizione del campione PQM (a.s. 2012/13)

Statistiche	ESCS medio	% ripetenti	% femmine	Dimensione classe
Min	-1.75	0	12.5	6
Media	-0.463	10.19	51.95	17
Max	0.71	60	100	31

## Statistiche per punteggio italiano del campione PQM (a.s. 2012/13)

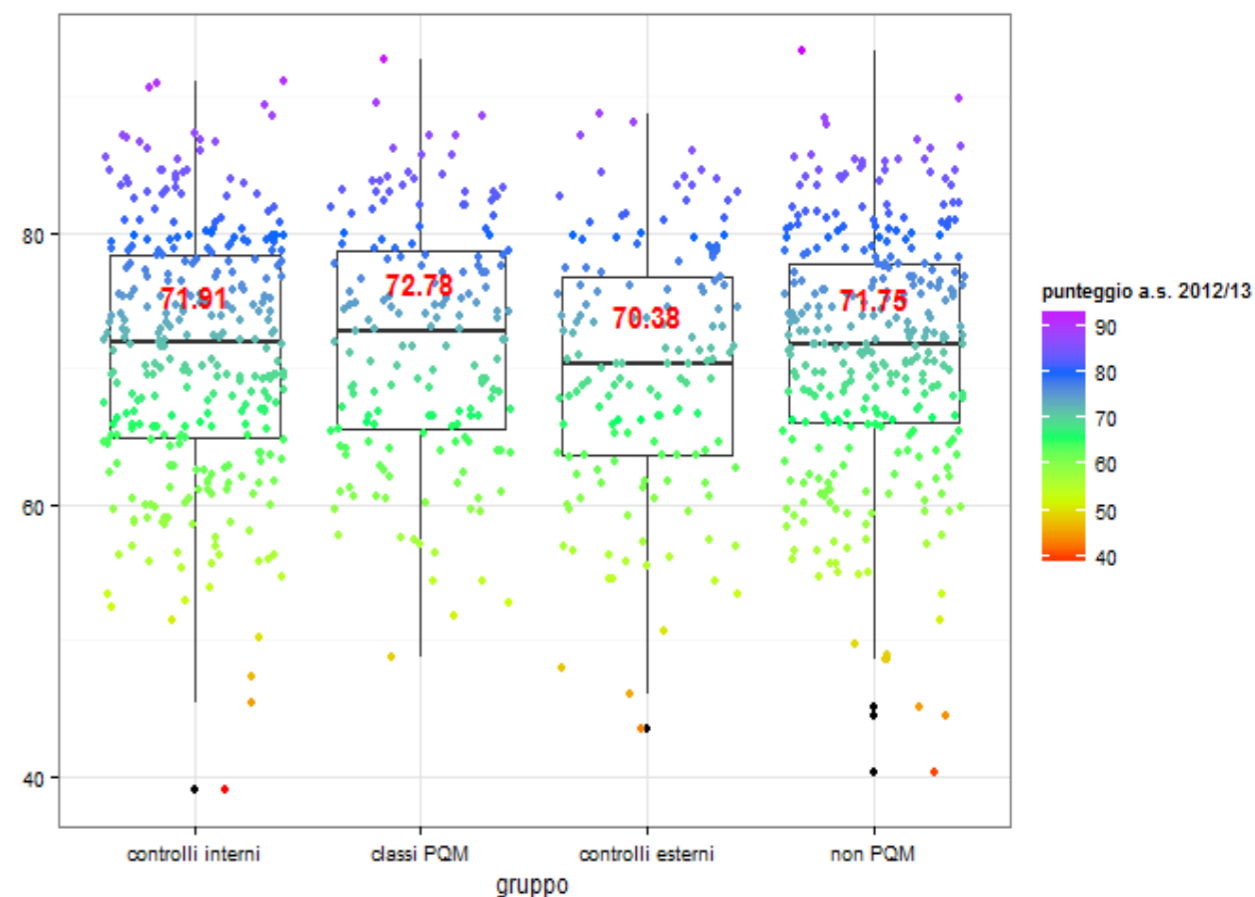
Gruppo	Q1	Media	Q3
Classi PQM	65.53	71.21 (70.66; 73.36)	78.68
Controlli interni	64.93	71.12 (70.12; 72.19)	78.27
Controlli esterni (psm)	63.70	69.80 (68.55; 71.60)	76.71
Non PQM	65.98	71.68 (70.30; 72.34)	77.68

IC al 95% per la media in parentesi

### EVIDENZA:

- Punteggio mediano classi PQM lievemente maggiore di quello delle classi non PQM (tutti i controlli).

## Distribuzione prestazioni italiano campione PQM (a.s. 2012/13)



# PQM e apprendimenti in italiano: focus sui processi

Le Linee guida sottese alla struttura e ai contenuti della **prova PQM** sono coerenti con i *QdR* e le Indicazioni Nazionali per la prova di italiano ([ubmiur.pubblica.istruzione.it/web/istruzione/](http://ubmiur.pubblica.istruzione.it/web/istruzione/))

Prova PQM composta da due parti:

- 1) **comprensione testo**: narrativo + espositivo;
- 2) **prova grammatica**

## ASPETTI VERIFICATI DA INVALSI:

• Nell'ambito della **comprensione del testo** si individuano **3 macro-processi**:

1. Comprendere e ricostruire il testo
2. Individuare le informazioni
3. Rielaborare il testo

La **prova di grammatica** presenta i seguenti **ambiti di indagine** (come da *QdR*):

1. Ortografia
2. Morfologia
3. Lessico e Semantica
4. Sintassi
5. Testualità
6. Formazione parole

## Comprensione della lettura: i macro-processi nelle prove PQM

### Comprendere e ricostruire il testo

- ✓ Comprendere il significato, letterale e figurato, di parole ed espressioni e riconoscere le relazioni tra parole.
- ✓ Fare inferenze dirette, ricavare informazioni implicite nel testo e/o tratte dell'enciclopedia personale del lettore.
- ✓ Cogliere relazioni di coesione e coerenza testuale (coesione logica entro e oltre la frase).
- ✓ Ricostruire significato di una parte più o meno estesa e poi globale del testo, integrando più informazioni e concetti anche formulando inferenze complesse.



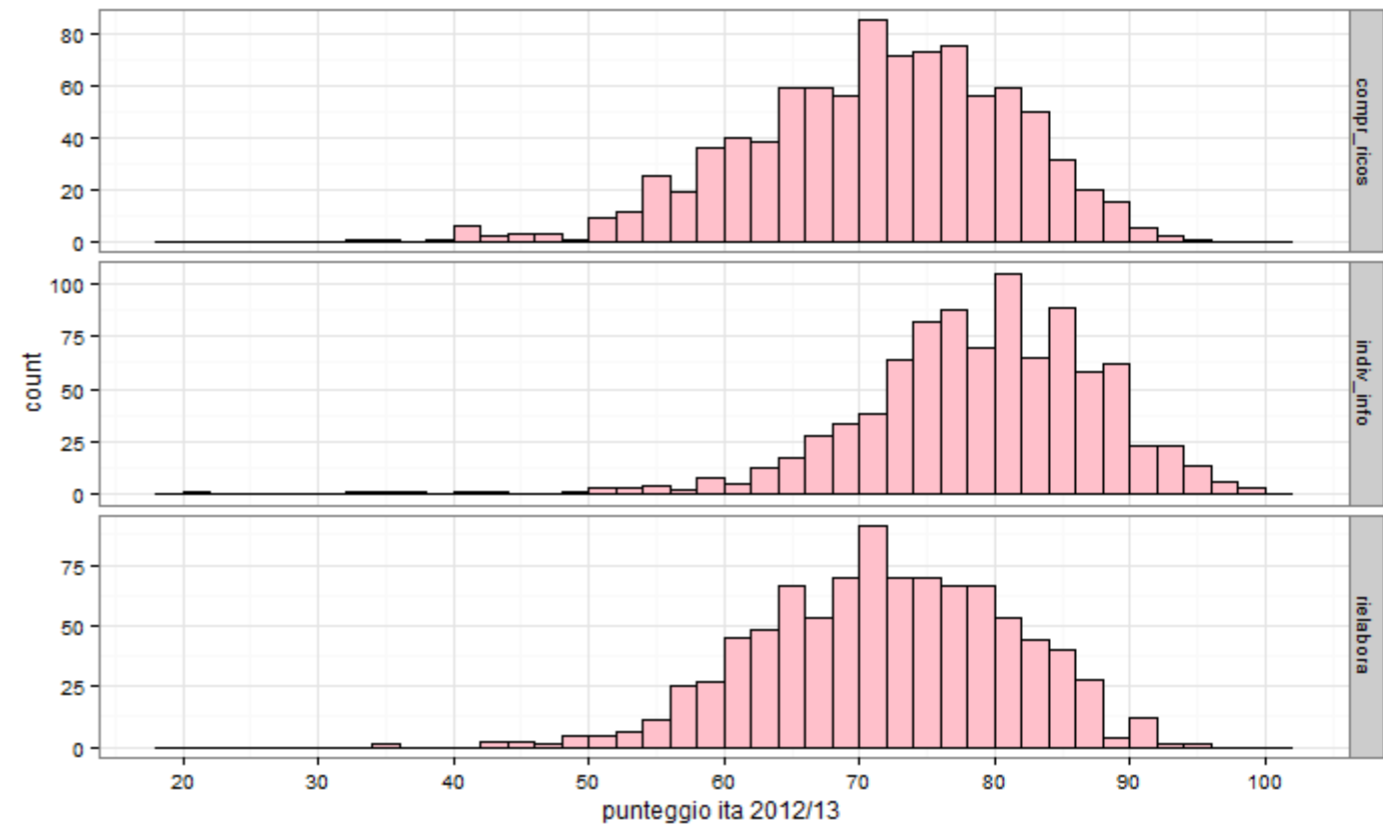
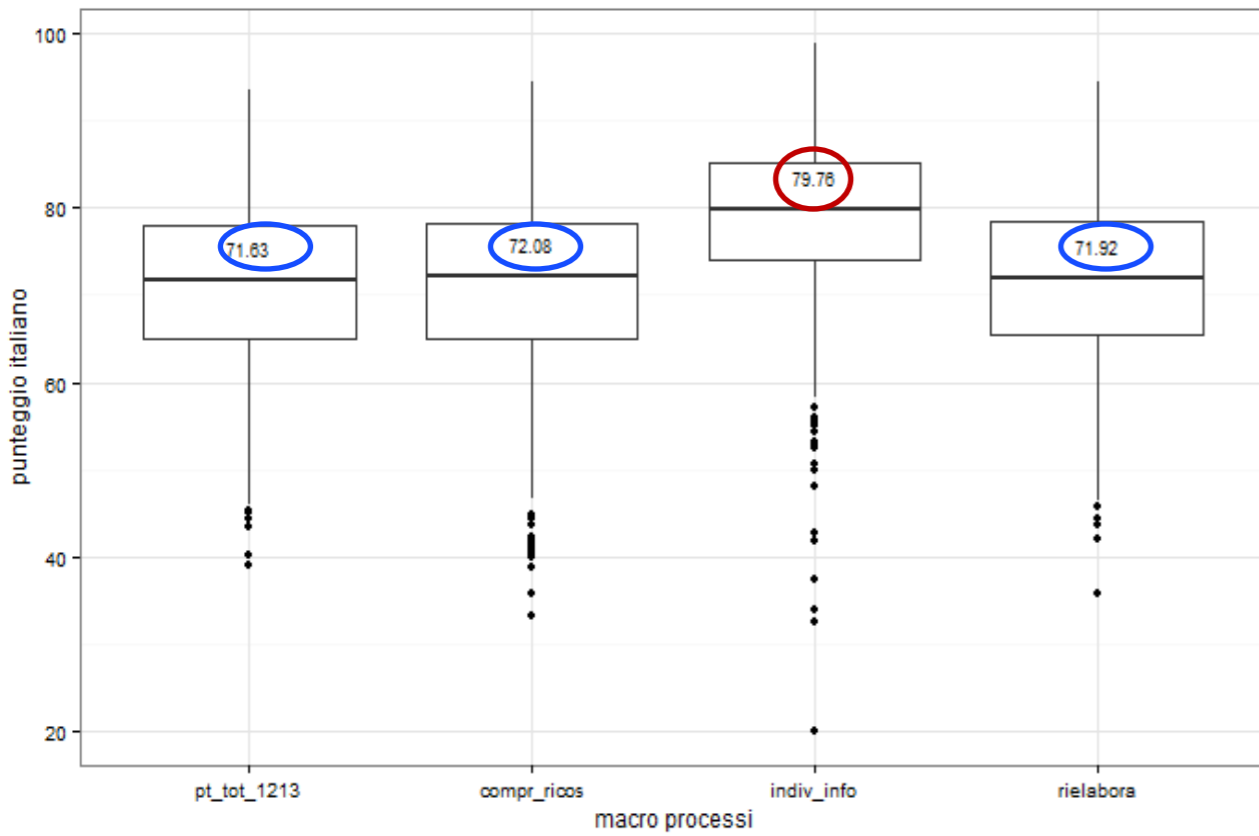
### Individuare Informazioni

- ✓ Individuare informazioni esplicitate nel testo.

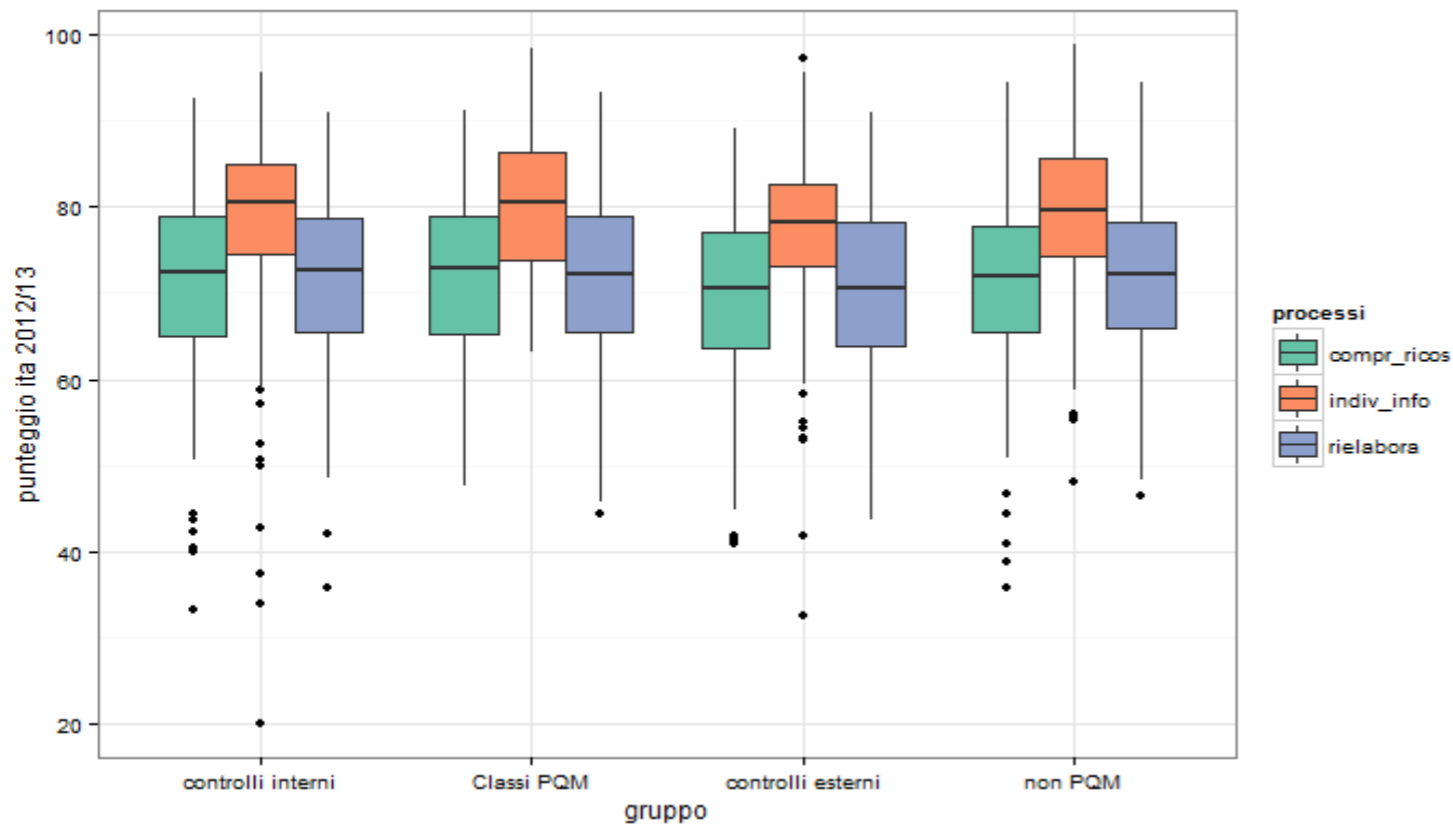
### Rielaborare il testo

- ✓ Sviluppare un'interpretazione del testo, partendo dal contenuto o dalla forma (al di là di una comprensione solo letterale).
- ✓ Riflettere sul testo e valutarne il contenuto e/ola forma alla luce delle conoscenze ed esperienze personali.

## Classi PQM a.s. 2012/13: distribuzione punteggi per macro-processi



## Distribuzione punteggi per macro-processi per gruppo

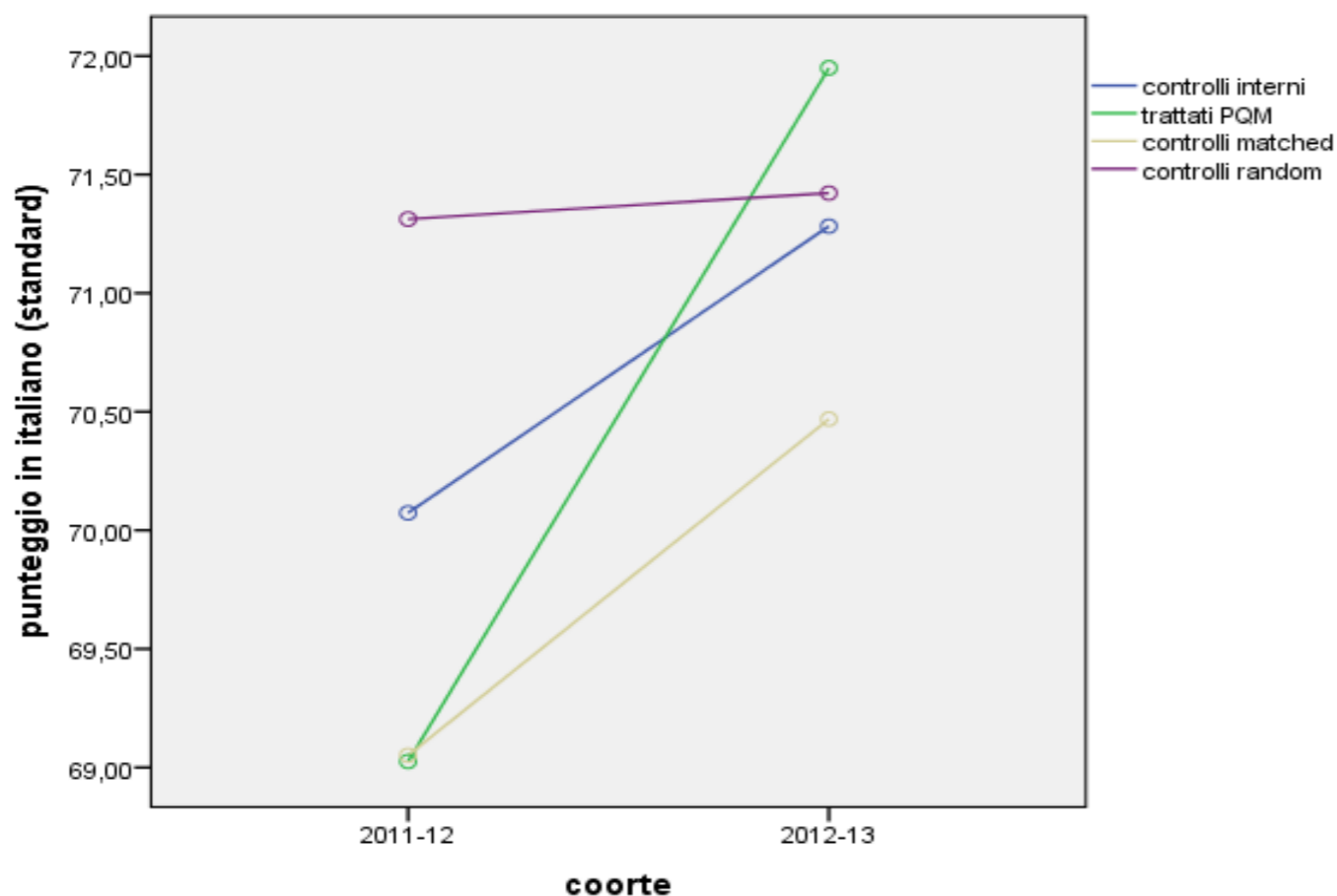


### ***Individuare Informazioni***

- Punteggio mediano più alto rispetto ai rimanenti e al punteggio della prova complessiva;
- *pattern* delle performance più alto per tutti i gruppi (trattati e controlli).

### ***Comprensione/Ricostruzione e Rielaborazione testo***

- Distribuzione performance più simmetriche rispetto ai valori medi e mediani;
- distribuzioni punteggi simili per trattati e controlli.



Parzializzazione per ESCS

## CONSIDERAZIONI:

- Differenza della prestazione tra partecipanti e non partecipanti varia in significativamente in funzione della coorte (p.value <.05).

- *Pre-intervento a.s. 2011/12*: per le classi PQM gap della performance a favore dei non trattati (in particolare, controlli interni e classi di scuole non PQM selezionate random).

- *Post-intervento a.s. 2012/13*: inversione tendenza. Prestazioni migliori per le classi PQM rispetto alle classi di controllo.

## Aspetti metodologici GLMM –Misure Ripetute (Gelman & Hill, 2007)

- Ipotesi di sfericità violata; si applica la correzione di Hyun e Feldt per i g.d.l.

- Normalità verificata;

- **Effetto netto del programma espresso dal termine di interazione tra trattamento e coorte (positivo e significativo con  $\alpha=5\%$ )**

- Parzializzazione rispetto ai valori dell'ESCS

Variabili	SSE (type III)	g.d.l.*	MSE	F	p. value
Coorte	543.942	1	543.942	10.475	0.001
Coorte*ESCS_m	0.348	1	0.348	0.007	0.935
<b>Coorte*Gruppo</b>	<b>423.405</b>	<b>3</b>	<b>141.135</b>	<b>2.178</b>	<b>0.044</b>
Errore(coorte)	46734.05	900	51.927		

## Media prestazioni italiano tra le coorti per gruppo

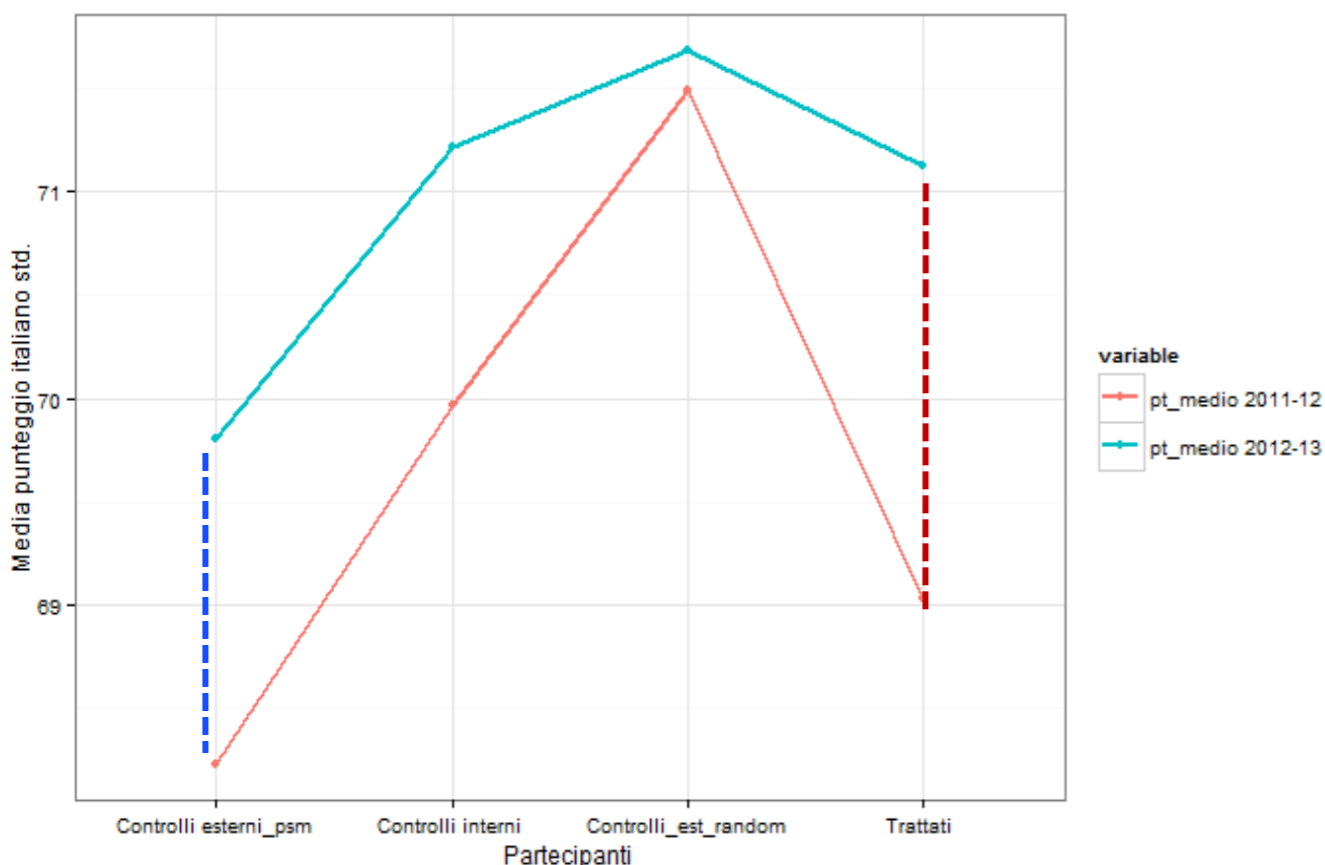
Gruppo	Media dei punteggi alla prova di italiano		Differenze (2011/12 – 2012/13)
	2011/12	2012/13	
Classi PQM	69.03	71.21	2.18
Classi non PQM (controlli esterni - psm)	68.22	69.8	1.58
Controlli interni	69.96	71.12	1.16
Non PQM	71.49	71.68	0.19
<b>Differenze (Trattati – Controlli esterni)</b>	0.81	1.32	<b>0.60</b>

### NOTA OPERATIVA:

Il confronto è in termini di punteggio medio in italiano per le classi PQM prima e dopo l'intervento.

Gap finale performance per gruppo tra le coorti

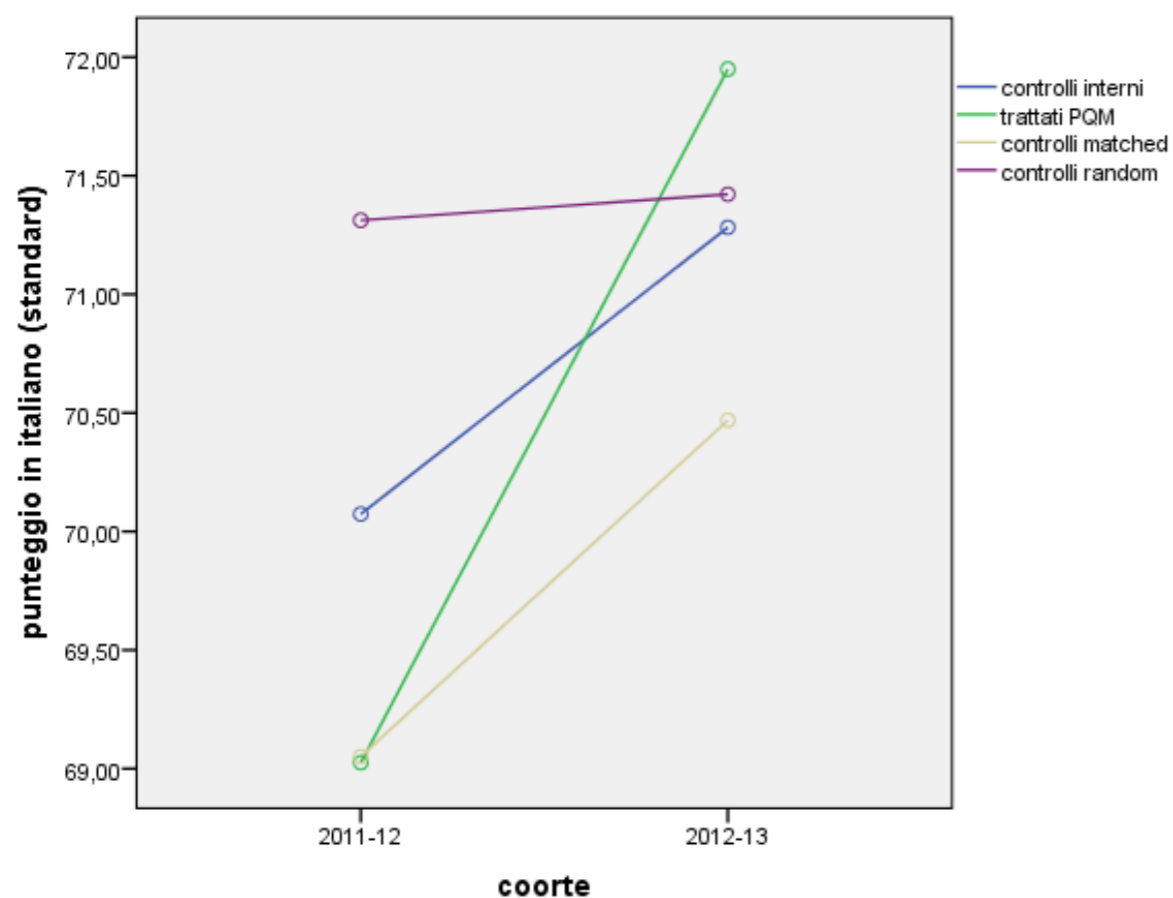
## Andamento medio prestazioni tra le coorti per gruppo



### EVIDENZA EMPIRICA:

Nel 2012/13 la differenza tra classi PQM e controlli è aumentata (1.32) a favore dei trattati rispetto alla differenza osservata nella coorte precedente (0.81).

**Questo vuol dire che PQM ha avuto impatto nella direzione attesa.**



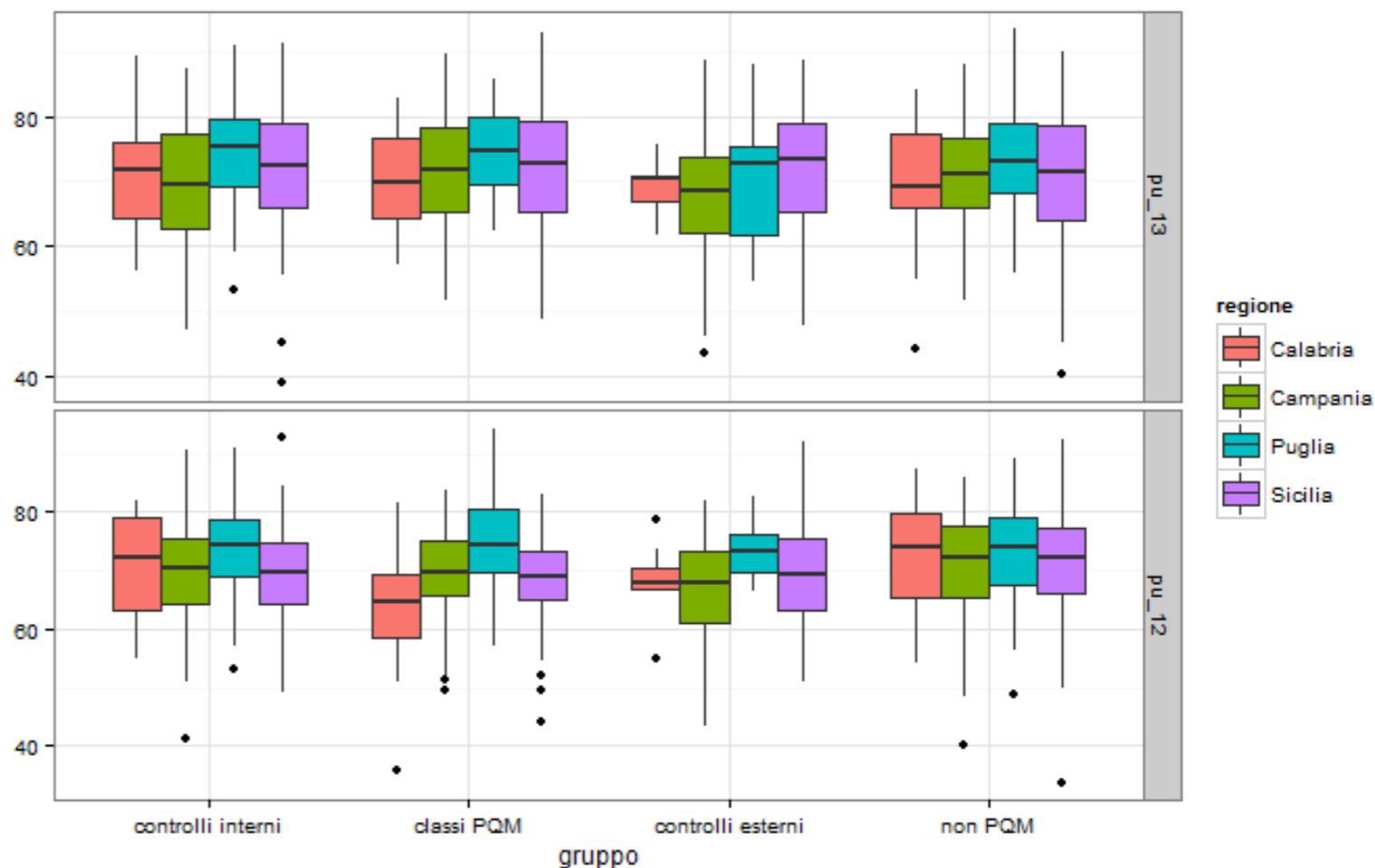
Parzializzazione per ESCS

## No effetto “spill over”

Nell’ambito delle classi non PQM, non si riscontrano differenze significative (p-value >0.05) nei livelli di prestazione tra **controlli interni** e “potenziali” PQM (**controlli esterni**) tra le coorti.

Gruppo	Media dei punteggi alla prova di italiano		Differenze (2011/12 – 2012/13)
	2011/12	2012/13	
Classi PQM	69.03	71.21	2.18
Controlli esterni (psm)	68.22	69.8	1.58
Controlli interni	69.96	71.12	1.16
Non PQM	71.49	71.68	0.19
<b>Differenze (Controlli esterni – Controlli interni)</b>	-1.74	-1.32	0.42

### Distribuzione punteggi italiano tra le coorti per gruppo



**Stima effetto “territorio” x trattamento:**

Evidenza riscontrata:

L’interazione nel modello è risultata non significativa (p-value>0.05).

### Livello medio delle prestazioni in italiano tra le coorti per gruppo

Regioni	Classi PQM		Controlli interni		Controlli esterni_psm		Classi non PQM	
	2011/12	2012/13	2011/12	2012/13	2011/12	2012/13	2011/12	2012/13
Calabria	62.90 (56.07; 69.71)	70.51 (65.42;75.58)	71.29 (66.42;76.15)	71.30 (66.46; 76.13)	64.84 (63.43;73.32)	70.80 (65.67; 72.19)	70.51 (68.59;75.68)	71.30 (66.04; 73.20)
Campania	69.31 (67.38;71.04)	71.03 (69.13;73.73)	69.31 (67.91;70.76)	69.88 (68.23;71.41)	66.76 (64.82;69.55)	67.04 (64.99; 70.45)	71.06 (68.58;71.94)	71.13 (69.78; 72.82)
Puglia	74.50 (69.79; 80.33)	71.46 (70.47;78.58)	73.15 (70.88;75.40)	74.07 (71.80;76.35)	73.35 (71.04;75.65)	69.97 (64.06;75.86)	72.86 (70.99; 75.23)	72.70 (71.08; 75.45)
Sicilia	68.78 (66.88;70.55)	71.61 (70.21;74.31)	69.37 (67.76;70.99)	71.51 (69.72; 73.28)	68.81 (67.10;71.24)	71.25 (69.75; 73.97)	71.37 (69.80;72.99)	72.31 (68.87; 72.78)



# Riflessioni conclusive

## La partecipazione al progetto PQM sembra migliorare significativamente gli apprendimenti in italiano

La differenza nei livelli di prestazione osservata tra le classi trattate e di controllo nel passaggio dal pre-intervento al post-intervento è significativamente attribuibile al programma (stime ottenute controllando rispetto ai valori dell'ESCS);

L'approccio empirico adottato è in linea con le tendenze nella letteratura prevalente sulla valutazione di politiche.

- *In particolare*, la metodologia di analisi contro-fattuale proposta ha permesso di sfruttare la ricchezza e i punti di forza dell'indagine nazionale sugli apprendimenti del Sistema Nazionale di Valutazione (SNV) per la costruzione del disegno di valutazione;

### Implicazioni:

PQM come modello “esportabile” per lo sviluppo e il miglioramento in termini di:

- ✓ cultura della valutazione
- ✓ qualità della ricerca sugli apprendimenti
- ✓ approcci innovativi alla didattica in italiano: non solo di quantità ma anche di *qualità* degli apprendimenti.



**Grazie per l'attenzione**