

## Un razionale statistico per l'uso del volume espiratorio forzato in 6 s\*

Robert L. Jensen, PhD; Robert O. Crapo, MD, FCCP; Paul Enright, MD; membri del Family Heart Study†

Lo scopo dello studio è stato quello di determinare il miglior surrogato dell'FVC quando si esegue una spirometria allo scopo di individuare la presenza di ostruzione o restrizione. Sono state analizzate le curve volume/tempo di 3.539 partecipanti del Family Heart Study con sessioni di controllo di qualità accettabili. Per ogni soggetto è stato determinato un indice della variabilità del volume emesso in un intervallo di tempo (FEV<sub>x</sub>) compreso tra 1 e 12 secondi. La variabilità minima registrata nell'ambito dello stesso test è stata osservata nel volume espiratorio forzato in 6 secondi (FEV<sub>6</sub>) e nel volume espiratorio forzato in 7 secondi (FEV<sub>7</sub>) (per entrambi il range medio è stato di 95 ml). La sensibilità e la specificità per diagnosticare l'ostruzione e la restrizione quando l'FVC veniva sostituito con il FEV<sub>6</sub> venivano quindi determinate prima e dopo aver tenuto conto di errori della misurazione pari a 50 ml (approssimativamente il limite inferiore della capacità degli spirometri di rilevare il volume). La sensibilità era del 76% prima dell'analisi dell'errore di 50 ml e del 95% dopo. La specificità era del 98% prima dell'analisi dell'errore di 50 ml e del 99,5% dopo. Concludiamo che l'uso del FEV<sub>6</sub> per sostituire l'FVC nei test spirometrici darà come risultato un miglioramento della riproducibilità, senza nessuna perdita significativa per la sensibilità o la specificità, dopo avere consentito l'introduzione di un errore di misurazione di 50 ml, per diagnosticare la presenza di ostruzione o di restrizione.

(CHEST Edizione Italiana 2007; 1:33-39)

**Parole chiave:** ostruzione delle vie aeree; FEV<sub>1</sub>; spirometria

**Abbreviazioni:** ATS = American Thoracic Society; IC = Intervallo di Confidenza; ERS = European Respiratory Society; FEV<sub>6</sub> = volume espiratorio forzato in 6 s; FEV<sub>7</sub> = volume espiratorio forzato in 7 s; FEV<sub>12</sub> = volume espiratorio forzato in 12 s; FEV<sub>x</sub> = volume a tempo; FHS Family Heart Study; LLN = limite inferiore del normale; NHANES III = Third National Health and Nutrition Examination Survey; TLC = capacità polmonare totale

**S**in dallo sviluppo della moderna spirometria, l'uso della capacità vitale, dell'FVC e del FEV<sub>1</sub> è stata una pratica standard. Un rapporto FEV<sub>1</sub>/FVC basso è l'attuale parametro standard per diagnosticare la

presenza di ostruzione delle vie aeree e il grado di riduzione del FEV<sub>1</sub> è impiegato per determinare la sua gravità.<sup>1</sup> Quando sono disponibili soltanto variabili spirometriche, l'FVC può essere impiegato per suggerire la presenza di restrizione.<sup>1</sup> Un'importante ragione al di là dell'utilità di tali parametri spirometrici è la loro riproducibilità intraindividuale, cosa che rende possibile calcolare i cambiamenti nel tempo e fare dei confronti tra i valori di riferimento con sicurezza.

Studi epidemiologici<sup>2-3</sup> hanno mostrato che le misure spirometriche sono correlate a numerosi eventi morbosi e comorbosi. Un comunicato del National Lung Health Education Program<sup>4</sup> ha proposto la diffusione della spirometria anche negli ambulatori di medicina di base al fine di individuare precocemente i casi di BPCO. Man mano che l'uso della

\*Dalle Pulmonary Divisions, LDS Hospital and University of Utah School of Medicine, Salt Lake City, USA.

†Il Family Heart Study è uno studio multicentrico stanziato dai National Institutes of Health.

Gli autori non sono coinvolti in alcun conflitto di interesse.

Manoscritto ricevuto il 18 gennaio 2006; revisione accettata il 15 giugno 2006.

La riproduzione di questo articolo è vietata in assenza di autorizzazione scritta dell'American College of Chest Physicians ([www.chestjournal.org/misc/reprints.shtml](http://www.chestjournal.org/misc/reprints.shtml)).

Corrispondenza: Robert L. Jensen, PhD, Pulmonary Division, LDS Hospital, Salt Lake City, UT 84143; e-mail: [Ldrjens1@ihc.com](mailto:Ldrjens1@ihc.com)

(CHEST 2006; 130:1650-1656)

spirometria si diffonde dagli ospedali o cliniche pneumologiche ad ambienti dotati di personale meno addestrato e con meno esperienza in materia pneumologica, è probabile che il livello di variabilità intraindividuale (“rumore”) aumenti, e ci si deve aspettare che vengano fatte delle classificazioni errate così come una minore abilità nell’individuare i cambiamenti nel tempo.

Una strategia finalizzata a minimizzare la percentuale di risultati falsi positivi è di impiegare parametri con la minore quantità di variazione intraindividuale. Quando il rumore nel segnale della spirometria è elevato e i valori misurati del soggetto sono molto vicini ai limiti minimi, il medico dovrebbe sentirsi meno sicuro su quale sia l’estremo della soglia di classificazione nella quale rientra il paziente. Parametri che hanno meno variabilità (rumore) permetteranno al medico di classificare con più fiducia valori vicini al limite. Si prenda in considerazione l’esempio di un soggetto con un  $FEV_1$  di 1,65 L che equivale al 77% del predetto. Se il limite inferiore di normalità (LLN) per questo soggetto era del 79%, una semplice classificazione lo etichetterebbe come anormale. Tuttavia, se il rumore nella misurazione era  $\pm 4\%$ , la classificazione non sarebbe così ovvia, mentre se il rumore era  $\pm 1,5\%$  la classificazione sarebbe più sicura.

Il National Lung Health Education Program<sup>4</sup> ha raccomandato la sostituzione dell’FVC con il volume espiratorio forzato in 6s ( $FEV_6$ ) per semplificare gli esami spirometrici. Probabilmente è stato scelto il  $FEV_6$  perché rappresentava il tempo espiratorio minimo raccomandato dall’American Thoracic Society (ATS) e perché erano disponibili i valori di riferimento. Abbiamo esaminato la variabilità intraindividuale nei tracciati spirometrici provenienti dal Family Heart Study (FHS) per determinare i livelli di tempo ai quali campionare il FEV ( $FEV_x$ ) e quali  $FEV$  potessero essere un buon surrogato per l’FVC. Dopo che i candidati sono stati scelti, la performance delle diverse grandezze candidate è stata paragonata al “gold standard” tradizionale per il  $FEV_1/FVC$  e FVC impiegati nella diagnosi dell’ostruzione e restrizione respiratoria attraverso la spirometria.

## MATERIALI E METODI

Sono stati sottoposti a spirometria 4.827 soggetti reclutati per l’FHS, uno studio multicentrico<sup>5</sup> comunitario finalizzato all’individuazione dei determinanti genetici e non-genetici della coronaropatia. I dati sono stati raccolti in quattro centri (Framingham, MA; Minneapolis, MN; Forsyth County, NC; e Salt Lake City, UT). I dati spirometrici sono stati ricavati da tecnici esperti usando gli standard dell’ATS del 1987.<sup>6</sup> Ogni partecipante ha effettuato da cinque a otto manovre spirometriche con uno spirometro a campana connesso ad un potenziometro lineare (Survey II; W.E. Collins; Braintree, MA). Il segnale del potenziome-

tro è stato analizzato attraverso un software installato su un personal computer (S&M; Quakertown, PA). Un sensore elettronico misurava la temperatura dello spirometro per ogni forma d’onda per consentire la correzione automatica dei valori misurati alla temperatura e pressione corporea in condizioni di saturazione.<sup>7</sup> Sia i dati grezzi della curva sia i dati calcolati venivano raccolti in floppy disk per la conservazione a lungo termine.

Copie dei floppy disk venivano spedite ad un centro di lettura (Salt Lake City) al fine di assegnare un punteggio alla qualità. Per ogni test sono stati assegnati da parte di due degli autori (R.L.J. o R.O.C.) punteggi di qualità totale simili ai punteggi accademici (da A ad F) per valutare il flusso e il volume. I punteggi relativi al volume sono stati assegnati dopo avere rivisto dati e tracciati sia dell’FVC che del  $FEV_1$ . I punteggi relativi al flusso sono stati assegnati dopo avere rivisto i tracciati del flusso del volume e i dati del picco di flusso. Il punteggio “F” stava ad indicare che lo studio era fallito secondo i criteri dell’ATS e che il recensore stabiliva che non erano stati ottenuti dei risultati sfruttabili. Il punteggio “D” indicava che lo studio non si conformava ai criteri dell’ATS, ma che il recensore giudicava alcuni dati sfruttabili. Per esempio, un test potrebbe avere uno sforzo iniziale buono ma un tempo di espirazione molto breve, così che il  $FEV_1$  sarebbe utilizzabile, mentre l’FVC non lo sarebbe. Il punteggio “C” indicava che i criteri dell’ATS per l’accettabilità e la riproducibilità erano stati pienamente soddisfatti. Punteggi del tipo “A” e “B” venivano assegnati in modo soggettivo e stavano ad indicare un miglioramento significativo nello sforzo, nella durata o nella riproducibilità al di sopra dei criteri minimi dell’ATS. I punteggi relativi alla qualità e i suggerimenti per un miglior rendimento successivamente venivano riportati ai tecnici e ai supervisori. L’attribuzione del punteggio ai dati è avvenuta durante la raccolta dell’FHS (1993-1995). Il presente studio è stato inizialmente proposto nel 1997 all’FHS come presentazione di un abstract e successivamente nel 1999 come manoscritto. I punteggi relativi alla qualità non sono stati rivisti e non vi era nozione del presente studio all’epoca dell’assegnazione del punteggio di qualità.

### *Dati per l’analisi della variabilità intraindividuale*

Al centro di lettura i tracciati grezzi e i dati computerizzati con i punteggi di qualità venivano trasferiti in un database per analizzarli. Il volume per ogni singola curva è stato misurato a 100 campioni al secondo (10 ms). È stato necessario soltanto misurare i dati ad intervalli di 250 ms dal momento che non sono stati osservati cambiamenti consequenziali nei parametri di variabilità tra i dati relativi a 10 ms adiacenti. L’analisi è iniziata a tempo zero (determinato dal volume retro-estrapolato) ed è continuata ogni 250 ms e successivamente fino a 12 s (Figura 1).

I dati venivano inclusi nello studio se il soggetto aveva almeno tre tracciati che si conformavano ai seguenti criteri: (1) tempo al picco di flusso < 250 ms (che dimostrava uno sforzo iniziale ragionevole); (2) tempo di espirazione  $\geq 10$  s (definito come tempo per l’FVC); e (3) valori dell’FVC compresi tra 1,0 e 9 l. Il limite di 10 s è stato stabilito in modo tale che l’indice di variabilità di ogni  $FEV_x$  potesse essere esaminato quando la maggior parte degli individui raggiunge il proprio FVC e tutti i soggetti hanno dati di  $FEV_6$ . Il tempo di espirazione di 10 s può limitare il set di dati perché alcuni individui non hanno un tempo di espirazione di 10 s. Tra 4.827 soggetti, 3.539 (73%) erano conformi ai criteri di inclusione. Ciò definiva il gruppo I e non restringeva i dati a confronto con i criteri di riproducibilità dell’ATS. Il gruppo I è stato definito in questo modo al fine di evitare che venisse introdotto qualsiasi bias sulla percentuale di variabilità. Le tre curve migliori per ogni individuo (definite dalla somma dei tre FVC più  $FEV_1$  più alti) sono state selezionate per l’analisi.

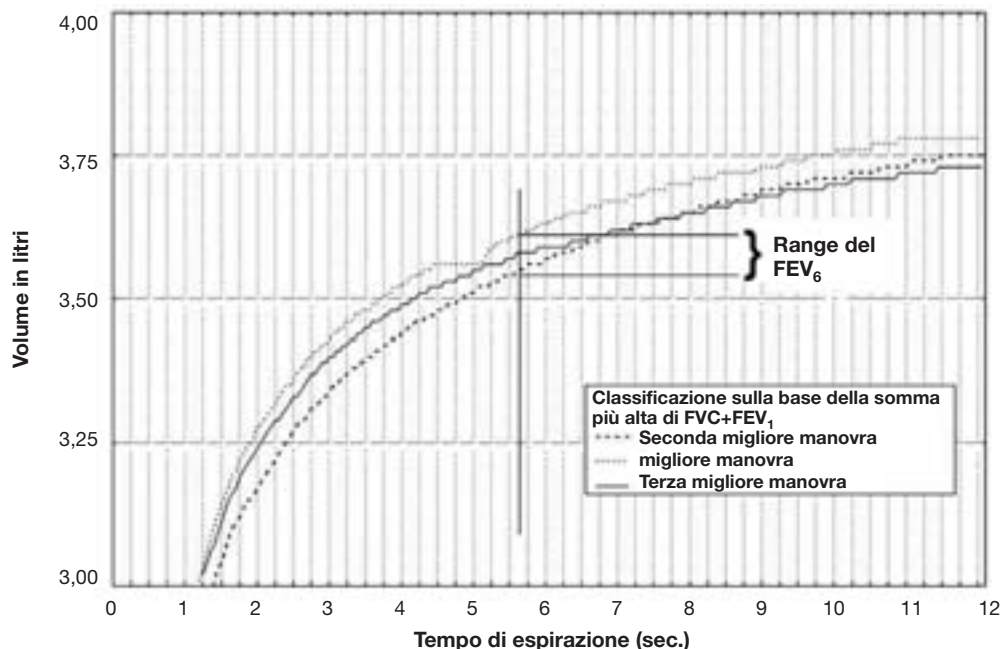


FIGURA 1. Esempio che mostra le tre manovre migliori di una sessione di test individuale, come stabilito dalla somma più alta dell'FVC e del FEV<sub>1</sub>. Per chiarezza, la porzione inferiore del grafico volume-tempo non viene mostrata. Si noti che la frequenza di campionamento è di 0,25 s e il metodo di calcolo per il range del FEV<sub>6</sub> è stato applicato anche agli altri candidati surrogati dell'FVC (FEV<sub>x</sub>) oggetto di valutazione.

Un secondo gruppo, il gruppo 2, è stato creato dal gruppo 1 con la richiesta addizionale che i test fossero conformi ai criteri di riproducibilità dell'ATS di  $\pm 200$  ml e che il loro punteggio di qualità del flusso e del volume fosse A, B o C. Il gruppo 2 conteneva 3.398 soggetti. Il gruppo 2 è stato creato in modo tale che le categorizzazioni dell'ostruzione sarebbero state affidabili e basate soltanto su buoni test di qualità.

#### Metodi statistici

**Analisi della riproducibilità:** I dati ricavati dal gruppo 1 sono stati impiegati per analizzare la variabilità durante il tempo di espirazione. Per ogni individuo, la variabilità è stata calcolata come il range di tre punti di volume ad ogni intervallo di tempo di 250 ms (range di FEV<sub>x</sub> dove x equivale al tempo). La Figura 1 mostra il calcolo per una unità di tempo per ogni soggetto. Soltanto i soggetti che avevano avuto almeno tre curve accettabili, e soltanto i dati ricavati dalle tre curve migliori sono stati impiegati per calcolare il range necessario allo studio. Tale operazione permetteva di eliminare e quindi di non includere nell'analisi gli sforzi spuri con livelli di FVCs molto bassi. Ogni 250 ms veniva tracciato il range medio del FEV<sub>x</sub> di tutti i soggetti in modo tale da illustrare la variabilità totale come funzione di tempo di espirazione fino a 12 s. Abbiamo scelto di analizzare i dati ricavati dal range in quanto gli standard dell'ATS e adesso dell'ATS/European Respiratory Society (ERS) enfatizzano il fatto che conformarsi alla ripetibilità significherebbe nei fatti conformarsi ad un criterio di differenza.<sup>8</sup> Ciò riflette anche il modo in cui la maggior parte di pneumologi, tecnici e produttori apprezzano la qualità di un test spirometrico e inoltre che gli errori di codifica e le segnalazioni d'errore del computer si basano sulle differenze.

Sono stati effettuati dei confronti statistici tra i range aggregati al FEV<sub>6</sub> e al volume espiratorio forzato in 7 s (FEV<sub>7</sub>), e tra il FEV<sub>6</sub> e il volume espiratorio forzato in 12 s (FEV<sub>12</sub>) utilizzando un t test per dati appaiati. L'errore di tipo I è stato corretto a

0,01, secondo la correzione di Bonferroni, per tenere conto dei tre confronti e per cautela nei confronti di qualunque conclusione statistica. Sono state calcolate anche le differenze medie intercorrenti nell'FVC tra FEV<sub>6</sub> e FEV<sub>7</sub>. I confronti statistici sono stati limitati a questi tre a causa dell'elevato numero delle potenziali combinazioni.

#### Analisi delle categorizzazioni diagnostiche

Relativamente ai soggetti del gruppo 2, ciascuno è stato classificato come normale, ostruito o ristretto facendo ricorso alle equazioni di riferimento del Third National Health and Nutrition Examination Survey (NHANES III).<sup>9</sup> Tali equazioni forniscono valori di predizione e i LLN per FVC, FEV<sub>6</sub>, FEV<sub>1</sub>/FVC e FEV<sub>1</sub>/FEV<sub>6</sub> che si basano su limiti statistici inferiori. Al momento il FEV<sub>6</sub> rappresenta il candidato surrogato dell'FVC per il quale i dati di riferimento sono disponibili. Sono stati scelti i dati di riferimento di Hankinson perché si basano su un campione a caso della popolazione generale degli Stati Uniti, le misure sono state effettuate con buoni controlli di qualità e le equazioni sono state fornite per tre differenti gruppi etnici.<sup>9</sup> I rapporti FEV<sub>1</sub>/FVC e FEV<sub>1</sub>/FEV<sub>6</sub> di ogni soggetto sono stati calcolati impiegando i FEV<sub>1</sub>, FEV<sub>6</sub> e FVC più alti ricavati da curve accettabili. Successivamente, ogni soggetto veniva classificato come ostruito o non ostruito facendo ricorso a questi due rapporti e ai LLN di Hankinson calcolati in base all'età, l'altezza, il sesso e la razza di ogni soggetto. L'ostruzione veniva considerata presente se anche ogni rapporto misurato risultava inferiore al limite più basso. La gravità dell'ostruzione veniva classificata in base alle raccomandazioni dell'ATS.<sup>1</sup>

In modo del tutto simile, veniva definito un modello restrittivo se si registravano FVC o FEV<sub>6</sub> bassi in assenza di ostruzione. Per ogni confronto, veniva costruita una Tabella 2 x 2. Venivano calcolati, sensibilità e specificità e i loro intervalli di confidenza al 95% (IC). Definire una restrizione basandosi sulla spirometria



FIGURA 2. Range medio di ogni intervallo di 250 ms calcolato da 3539 test nell'ambito delle tre manovre migliori (somme più alte di FVC + FEV<sub>1</sub>). Le barre di errore sono le DS dei range di ogni intervallo da 1 s. Il range medio minimo si verifica tra il FEV<sub>7</sub> e il flusso espiratorio forzato a 2,5 s.

comporta delle limitazioni.<sup>1</sup> L'assenza di un FVC ridotto o di un FEV<sub>6</sub> viene raramente associata ad una capacità polmonare totale (TLC) che sia al di sotto del range dell'LLN.

Le categorizzazioni di ostruzione venivano rianalizzate dopo avere introdotto la possibilità di un errore di  $\pm 50$  ml nel FEV<sub>1</sub> e nel FEV<sub>6</sub>, soltanto per il rapporto FEV<sub>1</sub>/FEV<sub>6</sub>. Cinquanta millilitri è un valore prossimo al limite in cui gli spirometri sono in grado di individuare una variazione di volume. Esso rappresenta approssimativamente dall'1 al 2% della maggior parte degli FVC ed è approssimativamente la metà della variabilità di FEV<sub>6</sub> osservato in questo studio. Per simulare l'effetto dell'introduzione dell'errore di 50 ml<sup>2</sup>, la seguente procedura è stata sottoposta ai dati di ciascun soggetto. Innanzi tutto, il rapporto FEV<sub>1</sub>/FEV<sub>6</sub> più grande possibile è stato calcolato aggiungendo 50 ml al FEV<sub>1</sub> e sottraendo 50 ml al FEV<sub>6</sub> (limite superiore). Secondo, il rapporto FEV<sub>1</sub>/FEV<sub>6</sub> più piccolo veniva calcolato sottraendo 50 ml dal FEV<sub>1</sub> e aggiungendo 50 ml al FEV<sub>6</sub>. Quando il FEV<sub>1</sub>/FVC calcolato andava al di sotto di questi limiti di rapporto FEV<sub>1</sub>/FEV<sub>6</sub> superiore e inferiore, le categorizzazioni cliniche venivano considerate equivalenti. Analogamente veniva introdotto un errore di 100 ml soltanto nel FEV<sub>6</sub> per effettuare un confronto tra i pattern restrittivi. Sensibilità e specificità venivano quindi ricalcolate e confrontate alle stime originali.

**Tabella 1—Dati della spirometria (n = 3539)\***

Variabili	(DS) Medie
FVC, L	3,93 (1,107)
FEV <sub>1</sub> , L	2,96 (0,899)
Picco di Flusso Espiratorio	7,68 (2,255)
FEV <sub>6</sub> , L	3,78 (1,084)

\*I più alti valori di FVC, FEV<sub>1</sub> di FEV<sub>6</sub> venivano ricavati dalle tre curve con le somme più alte FVC + FEV<sub>1</sub>. Il Picco di Flusso Espiratorio veniva preso dalla curva con la più alta somma FVC + FEV<sub>1</sub>.

## RISULTATI

Sono stati reclutati 4.827 partecipanti all'FHS. 3.539 soggetti sono stati inseriti nel gruppo 1 per analizzare la variabilità FEVx. Il gruppo 1 conteneva esami spirometrici per tutti gli individui che avevano completato almeno tre manovre con almeno 10 s di espirazione, e i dati venivano analizzati fino a 12 s. Alcuni di questi test non si conformavano ai criteri di riproducibilità dell'ATS di 200 ml (n = 141) ma venivano introdotti in modo tale che non vi fosse nessun errore sistematico nei calcoli di variabilità del FEVx. I dati spirometrici medi di questi soggetti sono stati riportati nella Tabella 1.

La variabilità media nel FEVx (espressa come il range medio) aumentava considerevolmente nel corso del primo mezzo secondo di espirazione, successivamente diminuiva lentamente fino approssimativamente a 7 secondi e infine aumentava alla fine dei dati analizzati a 12 s. Il range medio sia del FEV<sub>6</sub> che del FEV<sub>7</sub> era statisticamente inferiore al range medio del FEV<sub>12</sub> (p < 0,0001). Il range medio del FEV<sub>7</sub> è graficamente inferiore rispetto al FEV<sub>6</sub>; tuttavia, la differenza non è statisticamente significativa (p > 0,05).

Sulla base del FEV<sub>1</sub>/FVC, la prevalenza di ostruzione nel campione di 3.398 individui (il gruppo che era conforme ai criteri di accettabilità dell'ATS, con punteggi pari ad A, B o C) era del 15% (525 individui); di questi, il 68% aveva un'ostruzione media, il 24% aveva un'ostruzione moderata e l'8% aveva un'ostruzione grave. Utilizzando dei LLN<sup>9</sup> esatti, la sensibilità del FEV<sub>1</sub>/FEV<sub>6</sub> per diagnosticare un'ostruzione sulla base del rapporto FEV<sub>1</sub>/FVC era del

**Tabella 2—Diagnosi dell'ostruzione delle vie aeree con FEV<sub>1</sub>/FVC e FEV<sub>1</sub>/FEV<sub>6</sub> utilizzando i limiti inferiori esatti del range di riferimento e consentendo un errore di ± 50 ml nel FEV<sub>1</sub>/FEV<sub>6</sub>**

Variabili	FEV <sub>1</sub> /FEV <sub>6</sub> ostruzione	FEV <sub>1</sub> /FEV <sub>6</sub> assenza di ostruzione	Totali
Ricorrendo ai limiti minimi esatti del range di riferimento†			
FEV <sub>1</sub> /FEV <sub>6</sub> ostruzione	399	48	447
FEV <sub>1</sub> /FEV <sub>6</sub> assenza di ostruzione	126	2825	2951
Totale	525	2873	3398
Errori di classificazione, %	24	1,7	
Consentendo un errore di ± 50 ml nel FEV <sub>1</sub> /FEV <sub>6</sub> ‡			
Ostruzione FEV <sub>1</sub> /FVC	521	14	535
Assenza di ostruzione FEV <sub>1</sub> /FVC	4	2859	2863
Totale	525	2873	3398
Errore di classificazione, %	0,8	0,5	

\*I dati vengono presentati come N., a meno che non vi è altra indicazione.

†Sensibilità, 76,0% (IC, 72,1-79,5); specificità, 98,3% (IC, 97,8-98,8).

‡Sensibilità, 99,2% (IC, 97,9-99,8); specificità, 99,5% (IC, 99,2-99,7).

76%, e la specificità era del 98,3% (Tabella 2). La maggior parte dei risultati discrepanti era vicina al limite inferiore del range di riferimento. Dopo avere tenuto conto di un errore ipotetico di 50 ml, la sensibilità è aumentata fino al 99,2% e la specificità fino al 99,5% (Tabella 2).

Per il pattern spirometrico restrittivo, la sensibilità del FEV<sub>6</sub> era dell'88% e la specificità era del 98,8%. Quando veniva introdotto un errore potenziale di 100 ml nel FEV<sub>6</sub>, sensibilità e specificità aumentavano al 100% (Tabella 3).

#### DISCUSSIONE

Il confronto con i valori di riferimento necessita di una comparabilità sia biologica che tecnica tra la persona che viene sottoposta ai test e i valori di riferimento.<sup>10</sup> La riduzione della variabilità intraindividuale nelle misure spirometriche creerà un allineamento più vicino tra le popolazioni di riferimento e dovrebbe far diminuire gli errori nella classificazione clinica: innanzitutto, assicurando che ciascun valore misurato rappresenti più precisamente l'individuo che è stato sottoposto a test; secondo, posizionando più accuratamente le misurazioni dell'individuo al di sopra o al di sotto dell'LLN.

L'FVC ha una variabilità intrasoggettiva più alta del FEV<sub>6</sub>, in parte perché dipende dal tempo di espirazione. Tale problema è più evidente nei pazienti con ostruzione aerea, che non evidenziano un buon

plateau di volume neanche dopo 20 s di espirazione.<sup>11</sup> Espirazioni più lunghe si presentano anche in individui sani più anziani a causa della perdita del ritorno elastico. Man mano che il tempo di espirazione aumenta con l'età, e l'FVC diventa più dipendente dal tempo di espirazione, le possibilità di un errore di classificazione diagnostica aumentano. Per esempio, poniamo il caso che un soggetto di 65 anni venga sottoposto a test e lo esegua bene, espirando vigorosamente, ma soltanto per 6 s. Tale paziente si conforma ad uno degli standard ATS di fine test; tuttavia, le equazioni di riferimento per l'FVC si basano su dati ottenuti da soggetti con tempi di espirazione medi pari a 15 s. L'FVC del paziente sarà basso, relativamente al set di riferimento semplicemente perché il tempo di espirazione è più breve. Come risultato la proporzione tra FEV<sub>1</sub>/FVC misurato sarà alta relativamente al valore di riferimento ed è molto probabile che si verifichi un errore di classificazione come falso negativo.

Il FEV<sub>6</sub> è un potenziale surrogato dell'FVC. È approssimativamente dal 90 al 95% dell'FVC, si trova ad un tempo determinato e quindi è riproducibile, e dovrebbe migliorare l'accuratezza diagnostica migliorando la comparabilità tra i dati del soggetto e quelli di riferimento. Il tempo di espirazione del FEV<sub>6</sub> è breve abbastanza da essere raggiunto dalla maggior parte dei soggetti, è accettabilmente riproducibile e dà un buon risultato se paragonato al rapporto FEV<sub>1</sub>/FVC tradizionalmente impiegato nella categorizzazione dei pazienti. Il FEV<sub>6</sub> ha dei vantaggi come surrogato dell'FVC per le ragioni che seguono: (1) rappresenta il tempo minimo di espirazione definito accettabile sia dall'ATS che dall'ERS; (2) la variabilità intraindividuale nel FEV<sub>6</sub> è vicina al minimo della variabilità di tutti i FEV<sub>x</sub>, e non è significativamente differente dal FEV<sub>7</sub>, dove si realizza il valore minimo; e (3) dati di riferimento sono attualmente disponibili per il FEV<sub>6</sub> e il rapporto FEV<sub>1</sub>/FEV<sub>6</sub> ricavato dai dati NHANES III che non sono disponibili per altri candidati FEV<sub>x</sub>.

**Tabella 3—Diagnosi della restrizione spirometrica usando LLN precisi\***

Variabili	FVC Basso	FVC Normale	Totali
FEV <sub>6</sub> basso	242	30	272
FEV <sub>6</sub> normale	33	2.568	2.601
Totali	275	2.598	2.873

\*Restrizione FVC o FEV<sub>6</sub> bassi in assenza di ostruzione delle vie aeree. Sensibilità, 88% (IC, 83,4-91,5); specificità, 98,8% (IC, 98,3-99,2).

Durante questo studio non abbiamo riesaminato le curve spirometriche singole. Soltanto durante la fase iniziale della raccolta dati dell'FHS, gli autori (R.L.J. e R.O.C.) hanno esaminato visivamente le curve di ogni partecipante e ne hanno calcolato la qualità. Tale analisi si basa proprio su questi calcoli iniziali.

È stato dimostrato che impiegare le misure spirometriche per la diagnosi di restrizione è difficile. Quando l'FVC è normale, vi è una probabilità bassa che il TLC sia inferiore all'LLN<sup>12-13</sup>; tuttavia, se l'FVC è al di sotto dei limiti normali, la probabilità di un TLC clinicamente basso aumenta dal 55 al 57%. Anche se è una probabilità relativamente alta che il TLC possa essere basso, è necessario comunque che venga effettuato un test addizionale del volume polmonare per determinare la reale presenza di un vero pattern restrittivo.

Swanney e coll.<sup>14</sup> hanno pubblicato dei confronti effettuati tra 337 individui che si rivolgevano ad un laboratorio polmonare a Christchurch, NZ, per dei test clinici. La prevalenza di ostruzione nel suo campione era pari al 66%. La sensibilità e la specificità del FEV<sub>1</sub>/FEV<sub>6</sub> nel predire l'ostruzione aerea attraverso un rapporto FEV<sub>1</sub>/FEV<sub>6</sub> era del 96% e 97% rispettivamente. Quando veniva introdotto un errore possibile di ± 100 ml, sia la sensibilità che la specificità erano essenzialmente del 100%. Inoltre, è stato trovato che la variabilità intraindividuale nel rapporto FEV<sub>1</sub>/FEV<sub>6</sub> era dal 25 al 30% più bassa che nel rapporto FEV<sub>1</sub>/FVC.

Vi sono delle ragioni tecniche per scegliere l'FVC<sub>6</sub> come surrogato dell'FVC. Le espirazioni di 6 s sono più facili per i pazienti ed è meno probabile che li si affatichi. L'uso del FEV<sub>6</sub> consente una fine test ben definita contrariamente all'FVC. Diversi produttori hanno già inserito i calcoli del FEV<sub>6</sub> nei loro algoritmi.

Gli unici dati di riferimento per il FEV<sub>6</sub> disponibili sono quelli provenienti dai dati NHANES III dei residenti statunitensi e possono essere o non essere applicati ad altre popolazioni. Prima che il FEV<sub>6</sub> possa essere ampiamente applicato sono necessari i dati di riferimento provenienti da altre popolazioni o i confronti che stabiliscono la comparabilità con le equazioni NHANES III. Si stanno sviluppando le moderne equazioni di riferimento per l'Europa continentale, l'India e la Cina; tuttavia, finché non vengono pubblicate suggeriamo che le equazioni di predizione del NHANES III per il rapporto FEV<sub>1</sub>/FEV<sub>6</sub> siano un'alternativa ragionevole. Questo perché il rapporto FEV<sub>1</sub>/FEV<sub>6</sub> è in qualche modo insensibile all'altezza, al sesso, all'etnia e ricorre soltanto all'età per predire il rapporto. Per esempio, in un uomo di 40 anni, tali equazioni mostrano davvero poche differenze tra bianchi, razza afro-americana e ispanici o nel rapporto FEV<sub>1</sub>/FEV<sub>6</sub> predetto (range di 1,8%) o nei corrispondenti LLN per il rapporto FEV<sub>1</sub>/FEV<sub>6</sub> (range di 1,9%).

Un'ulteriore problematica sorge con un surrogato come il FEV<sub>6</sub>; in particolar modo, cosa fare se un soggetto non espira per 6 s. Tale argomentazione è particolarmente rilevante per gli individui giovani che solitamente espirano per meno di 6 s. Per i giovani che rientrano nel raggio d'età coperto dal NHANES III (≥ 8 anni), si può ricorrere tranquillamente alle equazioni NHANES III. Nel NHANES III venivano inclusi nell'analisi del FEV<sub>6</sub> soltanto FVC associati ad un plateau nella curva volume-tempo. Pertanto si può fare un confronto in termini di espirazioni complete brevi tra il FEV<sub>6</sub> misurato nei bambini e i valori di riferimento del FEV<sub>6</sub> nel NHANES III. Chiaramente, è necessario più lavoro per i bambini per definire la fine test. Per gli adulti che non espirano per 6 s, una potenziale soluzione sarebbe di predire il FEV<sub>6</sub> dalle curve con i tempi di esalazione più brevi. Ciò è stato dimostrato predire il FEV<sub>6</sub> nell'ambito di ± 40 ml ma è stato pubblicato soltanto in formato di abstract.<sup>15</sup>

L'analisi dell'errore è coerente secondo le raccomandazioni dell'ATS in base alle quali bisogna interpretare i valori misurati vicini alla soglia con cautela.<sup>1</sup> Vi è sempre una possibilità reale che un singolo soggetto possa essere al di sotto o superare abbondantemente l'LLN dopo avere rieffettuato un test o persino nell'ambito di un'unica sessione di test se i valori misurati risultano vicini all'LLN. L'analisi dell'errore in questo studio per le categorizzazioni diagnostiche è stata considerata conservatrice dal momento che 50 ml è soltanto un quarto dello standard di riproducibilità dell'ATS di ± 200 ml<sup>6</sup> e un terzo del nuovo standard di riproducibilità ATS/ERS di ± 150 ml.<sup>8</sup> Abbiamo usato inizialmente un'analisi d'errore di ± 200 ml. A tale livello, sensibilità e specificità erano del 100%, in accordo perfetto. Successivamente, abbiamo ridotto l'analisi dell'errore di ± 50 ml ed abbiamo applicato l'analisi dell'errore soltanto al rapporto FEV<sub>1</sub>/FEV<sub>6</sub>, lasciando invariato il rapporto FEV<sub>1</sub>/FVC. Si tratta di una modificazione pari a ± 2,2% per un rapporto FEV<sub>1</sub>/FEV<sub>6</sub> del 75%, basato su un FEV<sub>1</sub> di 3000 ml e un FEV<sub>6</sub> di 4000 ml.

Condurre un'analisi sugli effetti del rumore o sugli errori di misura è critico ai fini della comprensione della comparabilità dei nuovi metodi potenziali per classificare i pazienti con metodi clinicamente accettati per la classificazione diagnostica. Innanzitutto, gli LLN sono stimati soltanto da popolazioni sane ed hanno la loro variabilità e imprecisione. In uno studio<sup>16</sup> condotto su soggetti Indiani Americani, il 95-esimo percentile inferiore per l'FVC e il FEV<sub>1</sub> aveva dei IC di approssimativamente ± 2%. In secondo luogo, le attrezzature per la misura introducono un certo rumore. Nelle recenti valutazioni (2005) degli spirometri Collins Survey (W.E. Collins; Braintree, MA; lo stesso modello utilizzato nell'FHS), è stata trovata una ripetibilità media con un generatore di forme d'onda polmonari, che impiegava soltanto l'aria della stanza per eseguire i test, pari ad un valore

compreso tra l'1% e il 2% per il FEV<sub>1</sub> e approssimativamente all'1% per l'FVC. Tali variazioni sono state osservate persino con un segnale di input ripetibile fino ad entro < 1 ml. Il test effettuato sull'uomo aumenterà ulteriormente la variabilità nell'ambito dello stesso test che è stimata essere pari al 2,7% per l'FVC e al 3,3% per il FEV<sub>1</sub>.<sup>17</sup> Terzo, l'errore di 50 ml era approssimativamente la metà della ripetibilità nello studio per l'FVC e il FEV<sub>1</sub>. Soltanto piccole modificazioni comprese entro il rumore normale per un determinato individuo possono rendere conto delle differenze osservate nelle classificazioni tra il rapporto FEV<sub>1</sub>/FEV<sub>6</sub> e FEV<sub>1</sub>/FVC. Infine, nel gruppo 2 i soggetti in cui la differenza tra il FEV<sub>6</sub> più alto e più basso era < 50 ml, il 36,6% delle differenze corrispondenti nel FEV<sub>1</sub> andavano in direzione opposta e la dispersione attorno al FEV<sub>1</sub> medio era di ± 200 ml. In generale, il FEV<sub>1</sub> diminuisce man mano che il FEV<sub>6</sub> diminuisce, così la nostra scelta di un errore pari a ± 50 ml può mettere in discussione di più il rapporto FEV<sub>1</sub>/FEV<sub>6</sub> con la stessa classificazione impiegata rispetto all'uso del rapporto FEV<sub>1</sub>/FVC. Tuttavia, ciò dovrebbe verificarsi soltanto con valori estremamente vicini all'LLN.

Le classificazioni diagnostiche sono ovvie quando i dati misurati sono ben al di sopra o ben al di sotto dell'LLN. Quando i valori misurati sono vicini all'LLN, vi è incertezza nella classificazione. Vi erano 468 soggetti (13,2%) nel gruppo 2 che si trovavano in una regione borderline definita pari a ± 2,2% dell'LLN. Di questi, in 363 soggetti (77,6%) il FEV<sub>1</sub>/FEV<sub>6</sub> concordava con la classificazione basata sul FEV<sub>1</sub>/FVC. Dei 105 risultati che non concordavano con le classificazioni, tutti tranne 18 (0,51%) hanno cambiato classificazioni per poi accordarsi con lo standard FEV<sub>1</sub>/FVC con l'introduzione dell'errore di ± 50 ml.

Il presente studio evidenzia principalmente il fatto che la variabilità intraindividuale dei FEV<sub>x</sub> si minimizza intorno al FEV<sub>6</sub>, FEV<sub>7</sub>. Quindi, non soltanto il FEV<sub>6</sub> ma anche i parametri derivati dal FEV<sub>6</sub> avranno una variabilità inerente più piccola. Siamo d'accordo con Pedersen, che in un editoriale recente ha notato che non vi è un reale vantaggio nel FEV<sub>6</sub> ai fini della diagnosi di restrizione. Viene eseguita con risultati scarsi tanto quanto l'FVC e nella migliore delle ipotesi può soltanto aumentare la probabilità di un TLD al di sotto dell'LLN pari approssimativamente al 55%. Il nostro lavoro inoltre fornisce supporto ai risultati di Akpinar-Elci e coll. e di Vandevoorde e coll. che mostrano dei livelli accettabili di specificità e sensibilità per la diagnosi di ostruzione (i livelli di Akpinar-Elci e al. erano di 92% e 98%, rispettivamente, e, per Vandevoorde e coll., erano di 94% e 93%, rispettivamente). La variabilità più bassa trovata nel FEV<sub>6</sub> può portare a delle categorizzazioni dell'ostruzione più precise rispetto a quelle che si riscontrano attualmente utilizzando l'FVC.

## BIBLIOGRAFIA

- 1 Becklake M, Cropp RO. ATS Statement: lung function testing: selection of reference values and interpretative strategies. *Am Rev Respir Crit Care Med* 1991; 144:1202-1218
- 2 Higgins M, Keller JB, Wagenknecht LE, et al. Pulmonary function and cardiovascular risk factors in young black and white men and women: the CARDIA study. *Chest* 1991; 99:315-322
- 3 Kannel WB, D'Agostino RB, Silbeshatz H. Use of vital capacity for cardiac failure risk estimation in persons with coronary disease and left ventricular hypertrophy. *Am J Cardiol* 1996; 77:1155-1158
- 4 Ferguson GT, Enright PL, Buist AS, et al. Office spirometry for lung health assessment in adults: a consensus statement from the National Lung Health Education Program. *Chest* 2000; 117:1146-1161
- 5 Higgins M, Province M, Heiss G, et al. NHLBI Family Heart Study: objectives and design. *Am J Epidemiol* 1996; 143: 1219-1228
- 6 Gardner RM, Hankinson JL, Clausen JL, et al. Standardization of spirometry: 1987 update; official statement of the American Thoracic Society. *Am Rev Respir Dis* 1987; 136: 1285-1298
- 7 Johnson LR, Enright PL, Voelker HT, et al. Volume spirometers need internal temperature sensors. *Am J Respir Crit Care Med* 1994; 150:1575-1580
- 8 Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J* 2005; 26:319-338
- 9 Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999; 159:179-187
- 10 Solberg HE, Grasbeck R. Reference values. *Adv Clin Chem* 1989; 27:1-79
- 11 Stoller JK, Buist AS, Burrows B, et al. Quality control of spirometry testing in the registry for patients with severe alpha-1-antitrypsin deficiency. *Chest* 1997; 111:899-909
- 12 Swanney MP, Beckert LD, Frampton CM, et al. Validity of the American Thoracic Society and other spirometric algorithms using FVC and forced expiratory volume at 6 s for predicting a reduced total lung capacity. *Chest* 2004; 126: 1861-1866
- 13 Aaron SD, Dales RE, Cardinal P. How accurate is spirometry at predicting restrictive pulmonary impairment? *Chest* 1999; 115:869-873
- 14 Swanney MP, Jensen RL, Crichton DA, et al. FEV<sub>6</sub> is an acceptable surrogate for FVC in the spirometric diagnosis of airway obstruction and restriction. *Am Rev Respir Crit Care Med* 2000; 162:917-919
- 15 Steltner H, Vogel M, Sorichtner S, et al. Analysis of forced expired volume signals using multi-exponential functions. *Med Biol Eng Comput* 2001; 39:190-194
- 16 Cropp RO, Lockey J, Aldrich V, et al. Normal spirometric values in healthy American Indians. *J Occupational Med* 1988; 30:556-560
- 17 Kunzli N, Ackermann-Lieblich U, et al. Variability of FVC and FEV<sub>1</sub> due to technician, team, device and subject in an eight centre study: three quality control studies in SAPALDIA; Swiss Study on Air Pollution and Lung Disease in Adults. *Eur Respir J* 1995; 8:371-376
- 18 Pedersen OF. FEV<sub>6</sub>: a shortcut to spirometry? *Eur Respir J* 2006; 27:245-247
- 19 Akpinar-Elci M, Fedan KB, Enright PL. FEV<sub>6</sub> as a surrogate for FVC in detecting airways obstruction and restriction in the workplace. *Eur Respir J* 2006; 27:374-377
- 20 Vandevoorde J, Verbanck S, Schuermans D, et al. Obstructive and restrictive spirometric patterns: fixed cut-offs for FEV<sub>1</sub>/FEV<sub>6</sub> and FEV<sub>6</sub>. *Eur Respir J* 2006; 27:378-383