

Corrispondenza

All'Editor. In un Supplemento del *Giornale Italiano di Cardiologia*¹ è stato pubblicato il lavoro dell'Istituto Superiore di Sanità (ISS) sui risultati a breve termine dei bypass aortocoronarici in Italia. Si tratta evidentemente di uno studio importante, come tutti i lavori che riportano un database nazionale ed in particolare un database che tratta di risultati clinici.

Molte poche categorie mediche hanno le capacità organizzative e culturali per produrre documenti di questo tipo, a mia conoscenza in Italia questo è il primo ed unico rapporto di questo genere; va dato atto quindi alla categoria dei cardiocirurghi una serietà e una professionalità non comuni. Proprio per il fatto di essere questo il primo studio di outcome in un paese senza tradizioni di questo tipo, lo rende così importante e anche lo espone a possibilità di errori.

Se osserviamo ciò che è accaduto e accade negli altri paesi in termini di esperienze accumulate e discussioni nate sui database di outcome, si capisce come sia difficile ottenere dati certi e come sia ancora più difficile trattare tali dati, poiché la loro pubblicazione ha un enorme impatto nella pratica clinica.

Le correnti di pensiero sono le più varie, da chi pensa sia giusto tenere i risultati esclusivamente per gli addetti a meno di variazioni fuori ogni logica, a chi pensa di rendere pubblico tutto ciò che attiene ai risultati fino a pubblicare la mortalità non solo per singolo ospedale, ma anche per singolo chirurgo.

Quale Socio in regola con il pagamento delle quote dell'Associazione Nazionale Medici Cardiologi Ospedalieri (ANMCO) e della Società Italiana di Chirurgia Cardiaca (SICCH) penso sia naturale che scriva all'organo ufficiale delle due società, per esprimere alcune opinioni sull'argomento.

Lo stesso lavoro che è ora pubblicato sul *Giornale Italiano di Cardiologia*, in modo più succinto, è stato pubblicato sull'organo ufficiale della Società Europea di Chirurgia Cardiotoracica². Dal momento che si tratta

di un argomento importante e discusso, l'Editore ha creduto opportuno per il giornale ed i lettori pubblicare due editoriali di accompagnamento^{3,4}, scritti da persone che per la Società Europea avevano qualche titolo di conoscenza sull'argomento.

Il Consiglio Direttivo della SICCH, dopo la pubblicazione dei risultati sulla stampa nazionale, ha pubblicato a pagamento sugli stessi giornali una lettera di critica dello studio.

Questo studio è stato un grandissimo impegno per tutti noi che abbiamo partecipato e sicuramente per il gruppo dell'ISS che ha elaborato i dati, purtroppo per differenti ragioni in questo tipo di studio vi sono *bias* che riducono la valenza scientifica del lavoro.

L'ISS ha prodotto un elegante modello statistico anche nell'idea di elaborare un modello italiano simile all'EuroSCORE, sarebbe stato molto utile confrontare i due modelli per validare quello italiano e se possibile renderlo ancora più affidabile dell'EuroSCORE, che ben sappiamo non essere esente da critiche.

Purtroppo il lavoro presenta *bias* pesanti che cercherò di elencare nel modo più lineare possibile:

- nel luglio 2003, l'ISS organizzò una conferenza a Roma per pubblicare i dati preliminari dello studio, furono presentati 18 000 pazienti arruolati con una mortalità media del 3.7%, alla fine dello studio nel settembre 2004 la mortalità era del 2.6% su 36 000 pazienti arruolati. Si è verificato quel che tutti gli statistici conoscono, l'effetto registro sull'outcome, i dati preliminari determinano una distorsione dei dati successivi, per tale ragione tutti gli studi seri non pubblicano mai risultati preliminari a studio in corso;
- il periodo di osservazione non ha avuto un inizio comune e lo studio non si è concluso allo stesso momento per tutti i centri, che sono stati liberi di presentare i risultati secondo differenti periodi; vi sono centri che hanno inviato periodi di osservazione di 6 mesi contro altri con periodi di 32 mesi, in

letteratura si definisce un periodo accettabile di osservazione almeno di 24 mesi. Basterebbe solo questo aspetto per inficiare la validità dello studio. Alcuni centri sono entrati a studio concluso;

- i fattori di rischio utilizzati per “aggiustare” la mortalità variano moltissimo e sono diversi dai dati medi pubblicati dalla SICCH nel suo rapporto del 2003, e dai fattori di rischio europei che sono stati analizzati per costruire il sistema EuroSCORE. Per esempio:

- shock cardiogeno, è presente tra lo 0.2 e il 7%, il dato nei registri pubblicati presenta un’incidenza dello 0.5%, sarebbe interessante sapere quanti dei colleghi che leggono questo articolo inviano alla chirurgia il 7% dei loro pazienti in shock cardiogeno;

- malattia polmonare, nello studio in esame varia tra il 3.5 e il 35%, mentre la variazione normale è tra il 4 e il 7%;

- angina instabile, nei dati riportati nello studio varia tra il 2.4 e il 62%, in tutti i database internazionali il valore è del 12%;

- urgenza, è riportata con una frequenza fino al 14% nello studio dell’ISS, mentre nel database della Società di Chirurgia Cardiotoracica americana e dell’EuroSCORE varia tra l’1 e il 4%.

È evidente che se la mortalità è aggiustata in conformità a fattori di rischio così dispersi si possono determinare errori che vanificano anche il miglior modello statistico possibile. Si può obiettare che in ogni modo le medie dei differenti fattori di rischio, nel lavoro dell’ISS sono in linea con i database internazionali; il problema è che l’ISS ha prodotto una graduatoria mettendo a confronto i vari centri e quindi non descrivendo delle medie. Per rimanere molto pratici, un centro ha presentato un’incidenza di shock cardiogeno del 7.2%, con una mortalità reale del 3.3%; la mortalità è stata aggiustata all’1.85%. Un altro centro con una mortalità reale del 3.10% ha avuto una mortalità aggiustata dell’1.72% molto probabilmente perché questo centro aveva un’incidenza di angina instabile del 62.5%. San Donato nel periodo di osservazione (29 mesi) ha presentato una mortalità del 3.4% che è stata aggiustata al 4.3% (nel nostro centro l’incidenza di shock cardiogeno è stata un normalissimo 0.5%).

È evidente che vi è un problema di aderenza alle definizioni di fattori di rischio.

Questo lavoro produce una graduatoria unica nella letteratura mondiale e nella conclusione gli autori avrebbero dovuto dare la giusta importanza; infatti, gli 8 centri considerati a migliore prestazione sono centri a basso volume di interventi con una media di 273 casi di bypass per anno; solo uno ne produce 425. Ancora più interessante è che tutti i centri che hanno un’attività > 500 interventi/anno di bypass sono a livello medio e due a livello basso. Da questo studio emerge che il centro più performante esegue meno di 150 bypass/anno, il centro non ha pronto soccorso e chiude durante l’estate.

Questo tipo di risultato, avrebbe dovuto mettere in guardia i ricercatori sull’esistenza di gravi problemi nello studio, ma evidentemente ciò non è accaduto.

Nelle conclusioni gli autori osservano che la mortalità è scesa dal 2.8% del 2002 al 2.4% nel 2003-2004, tale differenza non è significativa tuttavia gli autori considerano tale differenza un importante miglioramento verso la mortalità considerata, il *gold standard* dello 0.9%, valore ottenuto dai risultati dei centri a miglior performance. È strano che non una parola sia spesa dai ricercatori dell’ISS sull’effetto registro. Questa è una caratteristica che tutti quelli che si occupano di database conoscono benissimo; il fatto di confrontare eventi negativi porta inevitabilmente a cercare in assoluta buona fede di evitare pazienti troppo gravi o di caratterizzare i parametri in modo più “pessimistico”. Gli statistici americani conoscono benissimo questa “tendenza” che definiscono “gaming”.

Cito un solo lavoro; l’editoriale pubblicato sul *Journal of the American College of Cardiology* nel 2005 dal titolo “The big chill: the deleterious effects of public reporting on access to health care for the sickest patients”⁵ che compara la mortalità per le angioplastiche coronariche dello stato di New York dove vi è un registro (obbligatorio con dati controllati da funzionari del dipartimento della salute) e i risultati dello stato del Michigan dove non esiste un registro pubblico, il “gaming” è chiaramente spiegato.

Per fare un esempio, prendiamo in considerazione il mio ospedale, San Donato, dove tutti gli interventi sono confrontati con il sistema EuroSCORE ed i dati raccolti sono certificati da un Ente esterno all’ospedale; se noi rifiutassimo l’intervento a quei pazienti che presentano un EuroSCORE > 40% la mortalità si abbasserebbe del 50%.

Personalmente sono convinto che i database siano assolutamente necessari, e tale convinzione è dimostrata dal fatto che San Donato con 2062 pazienti è il maggior fornitore di dati allo studio dell’ISS, al database della SICCH e tra i maggiori afferenti al database della Società Europea di Chirurgia Cardiotoracica. Ritengo che tali raccolte debbano essere obbligatorie, certificate e la pubblicazione dei dati debba essere condivisa con chi deontologicamente e legalmente è responsabile della cura dei pazienti, nello specifico i cardiocirurghi.

Dobbiamo chiederci a cosa deve servire un database; la risposta banale è per migliorare il tipo di assistenza che offriamo noi chirurghi ai nostri pazienti. Siamo sicuri che creando le condizioni per cui i chirurghi si sentano in qualche modo sul banco degli imputati si faccia un bene a chi ha bisogno di un intervento? Se da domani decidessimo che pazienti con EuroSCORE > 40% non debbano essere operati, a chi rendiamo servizio? Al paziente, all’amministrazione del nostro ospedale che vede diminuire i costi, al registro che dimostra di aver migliorato l’outcome, a noi chirurghi?

Lo sforzo al miglioramento deve essere sempre presente e il confronto è la base di questa crescita. Database condotti in modo ineccepibile sono irrinunciabili. I

risultati devono essere discussi nelle sedi competenti, come per esempio gli assessorati alla sanità, o l'ISS stesso, in modo da intraprendere azioni tali da migliorare sia il livello dell'elaborazione dei dati sia gli standard assistenziali.

Spero che questo studio dell'ISS continui, esista una curva di apprendimento per tutti, sappiamo che i lavori di outcome clinico e chirurgico in particolare sono molto complessi e questa era la prima volta dell'ISS in questo campo.

Le Società Scientifiche dovrebbero avere un ruolo di *partnership* più stretto ed essere coinvolte maggiormente in questo tipo di studio, a proposito come mai tra gli autori di questo studio non vi è nemmeno un cardiocirurgo?

Sicuramente sono stato prolisso, ma sono convinto che i giornali scientifici debbano servire a sviluppare un dibattito necessario per far crescere consapevolezza e responsabilità in tutti coloro che in qualche modo sono responsabili della salute dei cittadini di questo paese.

Lorenzo Menicanti

Dipartimento di Cardiocirurgia
Istituto Policlinico San Donato
San Donato Milanese (MI)

Bibliografia

1. Seccareccia F, D'Errigo P, Perucci CA, et al, e il Gruppo di Ricerca del Progetto BPAC. Il Progetto BPAC. Studio degli esiti a breve termine degli interventi di bypass aortocoronarico nelle cardiocirurgie italiane. Risultati finali. *G Ital Cardiol* 2006; 7 (Suppl 1-2): 5S-47S.
2. Seccareccia F, Peducci CA, D'Errigo P, et al, for the Research Group of the Italian CABG Outcome Study. The Italian CABG Outcome Study: short-term outcomes in patients with coronary artery bypass graft surgery. *Eur J Cardiothorac Surg* 2006; 29: 56-62.
3. Nashef SA. The Italian job on the CABG outcomes. *Eur J Cardiothorac Surg* 2006; 29: 62-3.
4. Menicanti L. Editorial comment. *Eur J Cardiothorac Surg* 2006; 29: 63-4.
5. Turi ZG. The big chill: the deleterious effects of public reporting on access to health care for the sickest patients. *J Am Coll Cardiol* 2005; 45: 1766-8.

Risposta. Molti dei problemi affrontati nella lettera del Dr. Menicanti sono già stati discussi con dettagliate argomentazioni e motivazioni nel corso di precedenti corrispondenze¹⁻³. Ci sembra tuttavia utile affrontare nuovamente in modo sintetico alcuni dei problemi scientifici sollevati.

Il controllo del confondimento rappresenta il razionale per l'applicazione dei modelli di *risk adjustment*. Per definizione, un "fattore di confondimento" deve essere associato all'esito in studio (in questo caso la mortalità), eterogeneamente distribuito tra le categorie di esposizione (nel nostro caso tra ospedali), non deve agire come modificatore di effetto, né rappresentare un

fattore intermedio tra l'esposizione e l'esito in studio (ad esempio, una complicanza clinica).

Come più volte esplicitato¹⁻⁴, nello studio BPAC è stato utilizzato un modello multivariato costruito direttamente sulla popolazione in studio, e non un modello esterno quale il modello EuroSCORE, in quanto l'obiettivo specifico dello studio non era di tipo predittivo, bensì esplicativo, vale a dire: identificare e controllare i possibili fattori di confondimento dell'associazione tra esposizione ad un fattore di rischio (ospedale) e la mortalità a 30 giorni da intervento di bypass aortocoronarico.

L'uso di una funzione di rischio costruita direttamente sulla popolazione in studio diventa quindi indispensabile al fine di identificare e controllare i veri confondenti dei confronti. Ad esempio, l'età è il più importante e conosciuto fattore di predizione della mortalità. Se però l'età fosse omogeneamente distribuita tra gli ospedali, potrebbe essere esclusa dal modello di aggiustamento del rischio, semplicemente perché non rappresenta un fattore di confondimento. Ciò che non varia non confonde.

Quanto al suggerimento di confrontare il modello utilizzato nello studio BPAC con quello EuroSCORE, possiamo anticipare che una nuova analisi, in via di pubblicazione, porta a risultati che non si discostano di molto da quelli prodotti con l'analisi dello studio BPAC già pubblicata.

Proprio in considerazione del fatto che un fattore di rischio, affinché possa essere definito confondente deve, tra le altre caratteristiche, essere eterogeneamente distribuito tra gli ospedali in studio, solo i fattori che mostrano variabilità, in questo caso tra ospedali, meritano di essere studiati e offerti al modello di aggiustamento del rischio. È il caso dello shock cardiogeno, della malattia polmonare, dell'angina instabile, ecc.

Per quanto riguarda il confronto dei dati elaborati nello studio BPAC con quelli pubblicati dalla SICCH nel rapporto del 2003 sarebbe necessario, per questi ultimi, conoscere consecutività e completezza dell'arruolamento. Prescindendo da alcune inapproprietezze di metodologia statistica, volendo forzare alcune definizioni, e prendendo le "medie \pm DS" delle frequenze come buone, la variabilità della prevalenza di angina instabile rilevata nello studio BPAC (2.4-62%) non differisce da quella del rapporto SICCH 2003, indicato come 26.0 ± 11.4 con mediana di 26.8. Infatti, in base alle più elementari nozioni di statistica, quando si riportano i dati in questo modo si assume che essi siano distribuiti normalmente, che il 95% delle osservazioni siano comprese tra quella media \pm 2 DS e la totalità delle osservazioni (precisamente il 99%) tra la media \pm 3 DS. Quanto presentato nel rapporto SICCH indicherebbe che il range di variazione dell'angina instabile in quella popolazione è 3.2-48.8% (se si considera il 95% delle osservazioni) o addirittura 0-60.2% (per la totalità delle osservazioni). Pertanto, quel 26.0 ± 11.4 riportato nel rapporto SICCH sembra più simile a quan-

to rilevato nello studio BPAC, che a quei dati internazionali citati nella lettera (“... in tutti i database internazionali il valore è del 12%”), soprattutto se si considera che la mediana riportata è di 26.8 e che, conseguentemente, il 50% delle osservazioni è superiore a quel valore.

Inoltre, come già riportato in precedenti documenti, la prevalenza di fattori come l'emergenza o lo shock cardiogeno rilevata nello studio BPAC è assolutamente consistente con quanto presente nella letteratura internazionale³. La sospetta e troppo elevata frequenza di alcune patologie quali il diabete o la malattia polmonare può essere spiegata dal fatto che la popolazione BPAC è mediamente più anziana e con una maggiore prevalenza di uomini rispetto alla popolazione EuroSCORE (19.3% della popolazione BPAC è al di sopra dei 75 anni e il 79.1% sono uomini; in EuroSCORE solo il 9.6% è al di sopra dei 75 anni e gli uomini sono il 72%). Essendo alcune patologie più frequenti nelle popolazioni maschili e più anziane, queste differenze non soltanto sono spiegabili, ma addirittura attese.

La relazione “outcome-volume di interventi” mostra comportamenti complessi che variano nel tempo e tale associazione può essere nulla per tecniche o procedure sanitarie ormai consolidate e routinariamente eseguite anche negli ospedali più piccoli, come nel caso della chirurgia coronarica. Purtroppo, principalmente nel Sud-Italia, oltre che sporadicamente nel Nord, alcuni ospedali a più alto volume di interventi mostrano la più alta mortalità dopo intervento di bypass aortocoronarico e questa osservazione resiste a qualunque analisi di sensibilità. Che alcuni piccoli centri possano presentare performance invidiabili, porta l'autore della lettera ad affermare incautamente “Questo tipo di risultato, avrebbe dovuto mettere in guardia i ricercatori sull'esistenza di gravi problemi nello studio, ma evidentemente ciò non è accaduto”, piuttosto che a porsi il ragionevole dubbio se tali osservazioni siano valide. In particolare, è doveroso puntualizzare che il “centro più performante”, come viene definito il centro 33, effettivamente esegue poco meno di 150 bypass isolati per anno ma, come confermato dall'analisi che stiamo attualmente conducendo sulle schede di dimissione ospedaliera per gli anni 2001-2003, presenta un'attività cardiocirurgia costante durante l'intero anno solare (compresa l'estate) e conferma in pieno i buoni risultati ottenuti con lo studio BPAC.

Per ciò che riguarda la durata dello studio, non esiste alcuna evidenza scientifica a supporto dell'affermazione che una determinata durata dell'osservazione consenta i confronti. L'unico vincolo è la “potenza” dei confronti. Inoltre, per gli ospedali a più alto volume di interventi, le stime medie effettuate su lunghi periodi di osservazione potrebbero addirittura mascherare possibili eterogeneità temporali negli esiti. Comunque, il periodo minimo di reclutamento (almeno 6 mesi consecutivi) è stato definito, in accordo con i cardiocirurghi rappresentati nel comitato scientifico, prima di proce-

dere all'analisi dei dati. Tra i 64 centri analizzati, un solo centro ha raccolto dati per 6 mesi consecutivi, mentre gli altri hanno raccolto dati per periodi compresi tra 10 e 32 mesi (mediana 21 mesi).

Relativamente all'annoso problema delle “graduatorie” gli autori, diversamente da quanto dichiarato nella lettera, non hanno voluto “affrettarsi ad affermare che non era corretto comparare i centri tra loro, ma che bisognava solamente comparare il centro in esame alla media” per difendere qualche irregolarità nei dati pubblicati. L'obiettivo era invece di esplicitare alcune limitazioni della procedura di “standardizzazione indiretta” che, per ovvie ragioni metodologiche, non consente il confronto tra strati ma solo il confronto tra ciascuno strato e la popolazione di riferimento. Con i metodi di standardizzazione indiretta il sillogismo aristotelico (se A è diverso da B e C è diverso da B, allora A è diverso da C) non si applica, se non verificando l'omogeneità della distribuzione dei confondenti tra gli strati che si vorrebbero confrontare.

In quanto all'“effetto registro” (citato in tutti gli articoli BPAC finora pubblicati), fenomeno certamente noto agli epidemiologi, non può essere confuso con l'altrettanto noto fenomeno del “surgeon's risk averse behavior”. L'analisi dei dati BPAC attraverso la “mortalità aggiustata” rappresenta proprio il più efficace sistema in grado di contrastare quella possibile propensione dei chirurghi che vengono a conoscenza dei propri risultati, a rifiutare i pazienti più gravi al fine di migliorare la propria casistica. Se, infatti, i risultati venissero presentati esclusivamente sotto forma di “mortalità grezza”, tale propensione non riuscirebbe ad essere evidenziata; di contro i sistemi di aggiustamento del rischio, proprio perché riescono a tener conto della tipologia e quindi della gravità preoperatoria dei pazienti (*case-mix*), tendono a produrre mortalità predette più basse della mortalità grezza quando la gravità dei pazienti è minore di quella della popolazione di riferimento e viceversa. A condizione, ovviamente, di utilizzare informazioni valide e riproducibili sui reali confondenti dei confronti.

L'effetto registro ha invece un altro significato, certamente più positivo: si tratta di un fenomeno scientificamente ben documentato^{5,6} per il quale il solo impianto di sistemi di sorveglianza, di per sé, migliora la qualità delle prestazioni sanitarie e quindi i risultati ottenuti sui pazienti. È quanto si evidenzia ormai da decenni, a partire dai registri per le malattie infettive fino ad arrivare a quelli per le malattie croniche (tumori o malattie cardiovascolari), e che nulla ha a che vedere con gli atteggiamenti discriminatori degli operatori sanitari.

Per quanto riguarda la richiesta di un maggior coinvolgimento delle Società Scientifiche in questo tipo di studi e alla “authorship” nelle pubblicazioni, è sufficiente richiamare la composizione del Gruppo di Ricerca del Progetto BPAC, a tutti gli effetti coautore dell'articolo in questione. La SICCH e alcuni eminenti cardiocirurghi sono presenti sia tra i componenti del

gruppo dedicato al “Disegno dello studio, coordinamento e analisi” che nel “Comitato Scientifico”. Come sancito nel protocollo dello studio, è poi proprio il Comitato Scientifico ad approvare i contenuti degli articoli e ad autorizzarne la pubblicazione, esattamente come è avvenuto per questo articolo.

Infine, per quanto riguarda la possibilità che la statistica sia scienza imperfetta, a prescindere dalla confusione tra modelli “predittivi” e modelli “esplicativi”, tutti i sistemi di osservazione devono minimizzare due errori, quello “sistematico”, che limita la validità delle misure e quello “casuale”, che compromette precisione/riproducibilità, e nessun sistema di osservazione, quindi nessuna conoscenza, almeno nella popolazione umana, può annullare questi errori. Dell’errore casuale si da conto con significatività e limiti di confidenza delle stime; di quello sistematico, con il controllo di qualità dei dati e, negli studi osservazionali, con i metodi di controllo del confondimento e con la valutazione dell’interazione. Di questi metodi si è già lungamente discusso e su questo terreno di dibattito scientifico siamo sempre disposti a discutere. Con un’unica certezza, per dirla con Popper, che in futuro certamente qualcuno “falsificherà” i nostri risultati.

Un’ultima considerazione: la discussione, anche molto critica, dei risultati degli studi fa parte della dinamica della ricerca scientifica, anzi ne è strumento prezioso. Ciò è ancor più vero negli studi di epidemiologia clinica sulla valutazione comparativa di efficacia degli interventi. Questi studi possono essere realizzati solo con il contributo e il confronto di diverse competenze professionali, cliniche ed epidemiologiche.

Ognuna con le proprie conoscenze ed i propri limiti. Ognuna riconoscendo competenze e valori delle altre. E dichiarando, se ci sono, i propri conflitti di interesse.

**Fulvia Seccareccia¹, Carlo Alberto Perucci²,
Paola D’Errigo¹, Danilo Fusco²**

¹*Centro Nazionale di Epidemiologia, Sorveglianza
e Promozione della Salute*

Istituto Superiore di Sanità, Roma

²*Dipartimento di Epidemiologia*

ASL RME, Roma

Bibliografia

1. Seccareccia F, Perucci CA, Fusco D, D’Errigo P. Reply to Biondi-Zoccai et al. *Eur J Cardiothorac Surg* 2006; 29: 856.
2. Seccareccia F, Perucci CA, D’Errigo P, Fusco D. Reply to Hekmat et al. *Eur J Cardiothorac Surg* 2006; 29: 857-8.
3. Seccareccia F, Perucci CA, D’Errigo P, Fusco D. Concerning the Editorial comment by Dr. Menicanti. *Eur J Cardiothorac Surg* 2006; 29: 858-9.
4. Seccareccia F, Perucci CA, D’Errigo P, et al, for the Research Group of the Italian CABG Outcome Study. The Italian CABG Outcome Study: short-term outcomes in patients with coronary artery bypass graft surgery. *Eur J Cardiothorac Surg* 2006; 29: 56-62.
5. Tu JV, Naylor CD. Coronary artery bypass mortality rates in Ontario. A Canadian approach to quality assurance in cardiac surgery. Steering Committee of the Provincial Adult Cardiac Care Network of Ontario. *Circulation* 1996; 94: 2429-33.
6. Hannan EL, Siu AL, Kumar D, Kilburn H Jr, Chassin MR. The decline in coronary artery bypass graft surgery mortality in New York State: the rule of surgeon volume. *JAMA* 1995; 273: 209-13.