

**EL RISC DE LA CIRURGIA
CORONÀRIA A CATALUNYA:
MÈTODES I USOS DE LA
SEVA AVALUACIÓ**

TESI DOCTORAL



Aida Ribera Solé, 2006

Departament de Ciències Experimentals i de la Salut
Programa de Doctorat en Ciències de la Salut i de la Vida
Universitat Pompeu Fabra (UPF)

EL RISC DE LA CIRURGIA CORONÀRIA A CATALUNYA: MÈTODES I
USOS DE LA SEVA AVALUACIÓ

Memòria presentada per Aida Ribera Solé per a optar al títol de Doctora,
elaborada a partir del treball realitzat sota la direcció del
Dr. Gaietà Permanyer Miralda,
a la Unitat d'Epidemiologia del Servei de Cardiologia
de l'Hospital Universitari Vall d'Hebron.

Programa de Doctorat de la Universitat Pompeu Fabra, bienni 2003-2004.

Barcelona, 2006

Als meus pares

ÍNDIX

	Pàg.
PRESENTACIÓ.....	7
ACRÒNIMS.....	11
RESUM.....	13
INTRODUCCIÓ.....	15
La cirurgia coronària.....	16
L'ajust del risc en l'avaluació de resultats.....	18
Importància d'ajustar pel risc.....	18
Les escales de risc.....	19
L'estudi CIRCORCA i la influència del tipus de gestió hospitalària.....	23
OBJECTIUS.....	25
MÈTODES I RESULTATS.....	27
Disseny i població d'estudi.....	27
Estratègia de recollida de dades i variables de l'estudi.....	27
Article 1.....	31
Article 2.....	43
Article 3.....	65
DISCUSSIÓ.....	75
Les escales de risc.....	76
El problema dels factors de risc rellevants.....	76
El problema de la població de referència i la seva influència en la validesa dels dos models de risc: EuroSCORE vs escala local.....	79
La possibilitat de recalibració.....	81
Influència del tipus de gestió.....	82
Millora dels resultats associada als canvis en el procés assistencial.....	85
Els resultats de la cirurgia coronària sense circulació extracorpòria.....	85
Limitacions.....	88
Implicacions pràctiques i necessitats de recerca.....	88
CONCLUSIONS.....	93
BIBLIOGRAFIA.....	95
DOCUMENTS ANNEXES	
1. Estudi CIRCORCA.....	101
2. Formulari de recollida de dades i definicions de les variables d'estudi....	123
3. L'EuroSCORE i el model de l'AATRM.....	139
4. Comentari editorial.....	143

PRESENTACIÓ

Aquesta tesi doctoral es presenta en forma de publicacions originals segons la normativa aprovada per la Comissió de Direcció de Doctorat del Departament de Ciències Experimentals i de la Salut de la Universitat Pompeu Fabra. El treball inclou, en aquest ordre, un resum estructurat, un capítol d'introducció, els objectius del treball, un capítol de mètodes i resultats que inclou els tres articles originals publicats, un capítol de discussió, un de conclusions i quatre apartats annexes amb material complementari.

Els tres articles centrals que es presenten s'emmarquen en una línia de recerca més àmplia sobre l'estudi dels resultats de la cirurgia coronària que es va iniciar fa més de 10 anys a la Unitat d'Epidemiologia del Servei de Cardiologia (UESC) de l'Hospital Universitari Vall d'Hebron. Les hipòtesis plantejades en aquest treball són una conseqüència directa d'un estudi previ i, per aquesta raó s'inclouen en un apartat annex els dos articles originals derivats d'aquest primer estudi.

Els articles originals que conformen el nucli d'aquesta tesi són els següents:

Ribera A, Ferreira-González I, Cascant P, Pons JMV, Permanyer-Miralda G por el grupo de investigadores del estudio ARCA. Evaluación de la mortalidad hospitalaria ajustada al riesgo de la cirugía coronaria en la sanidad pública catalana. Influencia del tipo de gestión del centro (estudio ARCA). Rev Esp Cardiol 2006;59:431-40

Ribera A, Ferreira-González I, Cascant P, Pons JMV, Permanyer-Miralda G for the ARCA study group investigators. The EuroSCORE and a local model consistently predicted coronary surgery mortality and showed complementary properties. J Clin Epidemiol (en premsa)

Ferreira-González I, **Ribera A**, Cascant P, Permanyer-Miralda G for the ARCA study group. Outcomes in off-pump versus on-pump coronary artery bypass grafting stratified by operative risk profile: an assessment using propensity score. Eur Heart J 2006;27:2473-80

Presentacions prèvies

Tant els resultats globals com els resultats parcials per centre s'han donat a conèixer de manera confidencial als investigadors de cada centre participant. A més, alguns resultats parcials d'aquest treball s'han presentat en els següents congressos i reunions científiques:

Ferreira I, Ribera A, Permanyer-Miralda C, Cascant P, Soriano N. Grupo de investigadores del estudio ARCA. Evaluación de los resultados de la cirugía coronaria ajustados al riesgo en centros de gestión pública y privada. Análisis de la mortalidad de los primeros 806 pacientes. Comunicació oral. Congrés de la Sociedad Española de Cardiología (octubre del 2003)

Ferreira I, Ribera A, Permanyer-Miralda C, Cascant P, Soriano N. Grupo de investigadores del estudio ARCA Determinantes del uso de cirugía coronaria sin circulación extracorpórea. Sus resultados en condiciones de práctica clínica real. Congrés de la Sociedad Española de Cardiología (octubre del 2003)

Ribera A, Permanyer-Miralda G, Ferreira I, Cascant P y grupo de investigadores del estudio ARCA. Evaluación de los resultados de la cirugía coronaria en Cataluña. La importancia de ajustar por el riesgo. Comunicació oral a la 'XXII Reunión Científica Anual de la Sociedad Española de Epidemiología' (Cáceres, 27-29 d'octubre del 2004)

Ferreira I, Ribera A, Cascant P y Permanyer-Miralda G. Resultados de la cirugía de by-pass coronario sin circulación extracorpórea según el riesgo basal. Comunicacion oral a la 'I Jornada Científica IRYSS. La investigación de resultados y servicios de salud en España' (Las Palmas de Gran Canaria, 1-5 de noviembre de 2005)

Ribera A, Ferreira I, Cascant P, Pons JMV y Permanyer G. ¿Cuál es la mejor estrategia para el ajuste del riesgo en la evaluación de resultados de la cirugía coronaria? Comparación entre una escala local y el EuroSCORE. Pòster a la 'I Jornada Científica IRYSS. La investigación de resultados y servicios de salud en España' (Las Palmas de Gran Canaria, 1-5 de noviembre de 2005)

Finançament

Aquest treball ha estat finançat gairebé íntegrament per l'Agència d'Avaluació de Tecnologia i Recerca Mèdiques (exp. 061/22/2000) i parcialment per la xarxa IRYSS de l'Institut de Salud Carlos III (exp. G03/202).

Agraïments

Aquesta tesi és el resultat del treball i la col·laboració de moltes persones a qui estic agraïda. De la majoria, els pacients, no en puc dir els noms perquè en son masses i perquè ho manaven les lleis de la confidencialitat. A tots els investigadors del grup de l'estudi ARCA dels diferents centres participants, als caps dels Serveis de Cirurgia Cardíaca i als investigadors que han fet tota la feina de recollir les dades. Els seus noms son en aquestes pàgines. A tots els companys i amics de la Unitat d'Epidemiologia i del Servei de Cardiologia de l'Hospital Vall d'Hebron, tant els que hi son ara com els que hi han passat i de qui he après coses.

Als amics. Al meu amic David i al meu germà Francesc per tota l'ajuda.

Als meus pares, Joan i Mercè.

Als qui han estat els meus mestres en els últims deu anys: Gaietà Permanyer i Purificació Cascant.

ACRÒNIMS

AATRM: Agència d'Avaluació de Tecnologia i Recerca Mèdiques

ARCA: Avaluació dels Resultats de la Cirurgia Coronària a Catalunya

CEC: Circulació extracorpòria

CIRCORCA: CIRurgia CORonària a CAtalunya

CMBD: Conjunt Mínim Bàsic de Dades

DASI: *Duke Activity Status Index*

IC: Interval de confiança

ICD-9: Versió 9 de l'Índex de Codis Diagnòstics

OR: *Odds Ratio*

QVRS: Qualitat de Vida Relacionada amb la Salut

STS: *Society of Thoracic Surgeons*

UESC: Unitat d'Epidemiologia del Servei de Cardiologia de l'Hospital Vall d'Hebron

ROC: *Receiver Operating Characteristics*

RESUM

Antecedents. L'estudi dels determinants del resultat de la cirurgia coronària i l'ajust del risc mitjançant instruments específics s'ha generalitzat en els últims anys. A Catalunya s'han avaluat els resultats de la cirurgia coronària des de l'any 1994 i a partir de diverses fonts es pot formular la hipòtesi de que els resultats de la cirurgia coronària en el nostre àmbit han millorat en els últims anys però que encara existeix una diferència segons el tipus de gestió hospitalària. La cirurgia sense circulació extracorpòria ha esdevingut cada vegada més habitual en els últims anys però encara hi ha controvèrsia sobre la seva efectivitat real pel fet que els resultats dels assaigs clínics i dels estudis observacionals son sovint contradictoris.

Objectius. 1) Avaluar la mortalitat hospitalària de la cirurgia de derivació aortocoronària, ajustada al risc preoperatori, en malalts de la sanitat pública operats en centres de gestió pública i privada. 2) Comparar el funcionament d'un escala local (l'escala de l'AATRM) i l'EuroSCORE en la predicció del risc de mortalitat després de la cirurgia coronària. 3) Comparar l'efectivitat de la cirurgia sense circulació extracorpòria respecte de la cirurgia amb circulació extracorpòria segons el risc preoperatori.

Mètodes. Es van incloure de forma prospectiva tots els malalts consecutius sotmesos a una primera intervenció d'empelt coronari aïllat, programats entre octubre de 2001 i octubre de 2003, en tres hospitals de gestió pública i els dos hospitals de gestió privada que operen malalts procedents de la sanitat pública. El risc preoperatori es va avaluar mitjançant l'EuroSCORE i el model predictiu desenvolupat per l'Agència d'Avaluació de Tecnologia i Recerca Mèdiques. Es va comparar la validesa d'ambdues escales mitjançant indicadors de "validesa estadística" (discriminació i calibració) i de "capacitat avaluativa" (la influència del model en el càlcul de la mortalitat ajustada en els hospitals individuals). Per a l'estudi del resultat de la cirurgia sense circulació extracorpòria es va estratificar la població en dos grups segons el risc preoperatori (EuroSCORE >3 o ≤3) i es va calcular l'índex de propensió (probabilitat de ser operat mitjançant cirurgia sense CEC) global i per a cada estrat de risc, a partir de totes les variables clínicament rellevants. És va estimar el benefici de la cirurgia sense CEC en cada grup de risc ajustant per l'índex de propensió.

Resultats. Es van operar 1.605 pacients (el 21% en hospitals privats). Els centres privats tenien major proporció de pacients amb indicació no electiva (64% vs 50), amb angina inestable (17% vs 11) i amb grau funcional IV (20% vs 11). La OR per mortalitat hospitalària en centres privats vs centres públics ajustada per l'EuroSCORE va ser de 0,56 (IC95%: 0,29 – 1,06), ajustada per l'escala de l'AATRM de 0,56 (IC95%: 0,29 – 1,07). La mortalitat observada (3,55% en centres privats; IC95%: 1,7 – 5,5; i 5,13% en centres públics; IC95%: 3,9 – 6,3) no era significativament diferent a l'esperada segons

ambdós models. En fer l'ajust per edat, funció ventricular esquerra, estat preoperatori crític, disfunció renal, infart de miocardi recent, malaltia pulmonar obstructiva crònica, arteriopatia perifèrica i angina inestable, el tipus de gestió del centre mostrava una associació estadísticament significativa amb la mortalitat hospitalària: major supervivència en els hospitals privats amb una OR de 0,43 (IC95%: 0,21 – 0,87).

Comparant els resultats en els hospitals que havien participat també en un estudi previ es va observar que, mentre que la prevalença de la majoria d'indicadors de risc gairebé no ha canviat en els últims anys (p.e. mitjana d'edat de 65 anys vs 63, taxa d'intervencions urgents del 53% vs 55, malaltia pulmonar obstructiva crònica, 15% vs 16), la mortalitat hospitalària crua ha disminuït del 9% al 4,7%.

L'àrea sota la corba ROC era d'aproximadament 0,75 per ambdós models i la calibració era baixa, observant-se que l'escala local sobreestimava el risc de manera significativa. L'ordre dels hospitals segons la mortalitat ajustada era el mateix utilitzant ambdues escales però els centres qualificats com a *outliers* diferien segons el model utilitzat i abans i després de recalibrar les escales.

Els pacients operats sense CEC tenien menys risc de mortalitat postoperatoria (OR=0,53; IC95%: 0,3 – 0,94), d'infart postoperatori (OR=0,53; IC95%: 0,33 – 0,84), de reintervenció (OR=0,4; IC95%: 0,2 – 0,8) i d'estada postoperatoria de més de 10 dies (OR=0,55; IC95%: 0,4 – 0,7); essent aquest benefici major en els pacients de baix risc.

Conclusions. En els malalts operats actualment d'una primera derivació aortocoronària a Catalunya: 1) El tipus de gestió del centre s'associa amb una diferència marginal en el resultat ajustat pel risc, en el sentit de millor supervivència en els centres de gestió privada. 2) La comparació amb resultats previs i l'ús d'un instrument d'àmbit local han permès evidenciar una important tendència a la millora de l'efectivitat de la cirurgia coronària en els últims anys en ambdós grups de centres. 3) Els instruments d'ajust del risc són una eina útil per a l'avaluació del resultat de la cirurgia coronària. 4) En condicions de pràctica clínica normal la cirurgia sense circulació extracorpòria s'associa a menys esdeveniments desfavorables que la cirurgia amb circulació extracorpòria, essent aquest benefici major en els pacients amb risc preoperatori baix.

INTRODUCCIÓ

L'interès per avaluar els resultats de la cirurgia coronària a Catalunya no és nou. El primer estudi multicèntric sobre resultats quirúrgics a curt termini de la cirurgia cardíaca (coronària i valvular) realitzat a Catalunya es va dur a terme durant l'any 1994 en un estudi promogut per l'agència d'Avaluació de Tecnologia i Recerca Mèdiques (Pons JMV et al, 1997). Per primer cop en el nostre àmbit es van poder conèixer els resultats quirúrgics utilitzant sistemes d'ajust del risc.

De forma gairebé contemporània a la publicació dels resultats d'aquest estudi començà un altre estudi multicèntric (estudi CIRCORCA) promogut per la Unitat d'Epidemiologia del Servei de Cardiologia (UESC) de l'Hospital Vall d'Hebron i finançat per la mateixa Agència d'Avaluació de Tecnologia i Recerca Mèdiques, amb l'objectiu d'avaluar els resultats quirúrgics i la qualitat de vida relacionada amb la salut dels malalts sotmesos a cirurgia de revascularització coronària (Permanyer-Miralda G et al, 2001). Tant l'estudi de l'AATRM com el de la UESC pretenien entre d'altres coses estudiar la influència sobre el resultat de factors no clínics com és el tipus de gestió pública o privada del centre (Pons JMV et al, 1999) o el tipus de finançament públic o privat de l'atenció mèdica (Permanyer-Miralda G et al, 1998).

Des d'aquests dos estudis pioners s'han produït canvis importants: 1) La pràctica terapèutica de revascularització coronària al nostre país igual que a la resta del món ha evolucionat. 2) L'avaluació dels seus resultats ha esdevingut un exercici habitual, necessari i plenament acceptat per la comunitat científica i mèdica, i es troba encara en ple desenvolupament. 3) Això ha conduït globalment a un enriquiment del concepte de risc, augmentant la riquesa i precisió de les decisions clíniques.

Paral·lelament i com a conseqüència d'aquests canvis, s'ha generalitzat l'ús d'instruments d'ajust del risc estandarditzats. Gairebé de forma contemporània a la publicació dels resultats de l'estudi CIRCORCA es va publicar un model desenvolupat específicament per a ser utilitzat en el context europeu (Nashef SAM et al, 1999) que després ha estat promogut per les societats científiques de cardiologia i cirurgia cardíaca. Aquesta proliferació d'instruments desenvolupats en diferents àmbits ha estimulat la polèmica sobre quins instruments de mesura del risc són més apropiats en cada cas.

Aquests fets apunten la necessitat de conèixer quina és la situació actual de la cirurgia de revascularització coronària al nostre país i si els instruments teòricament més aplicables al context de la sanitat catalana són o segueixen essent vàlids en aquesta nova situació.

Per tal de comprendre plenament el problema que justifica el treball d'aquesta tesi estructurarem la introducció de la següent manera: en primer lloc, descriurem de manera general l'estat actual sobre la recerca de resultats de la cirurgia de revascularització coronària; en segon lloc, explicarem la importància d'ajustar pel risc i descriurem les eines que tenim per fer-ho; i finalment, resumirem breument els principals resultats de l'estudi CIRCORCA i les observacions que ens van portar a dissenyar l'estudi ARCA.

La cirurgia coronària

L'eficàcia de la cirurgia de revascularització coronària ha quedat ben establerta en assaigs clínics des de fa ja més de tres dècades (Yusuf S et al, 1974; BARI Trial Investigators, 1997). En un principi, es va comprovar el seu avantatge sobre el tractament mèdic en reduir la mortalitat i alleujar els símptomes en pacients de més risc i posteriorment es va comparar la seva eficàcia amb les tècniques de revascularització percutània. Actualment s'accepta que és més eficaç que la revascularització percutània en pacients amb afectació del tronc comú, malaltia de tres vasos, o amb malaltia difosa, disfunció ventricular severa, i en els malalts diabètics (Rihal CS et al, 2003; BARI Trial Investigators, 2000).

Els estudis observacionals sobre grans registres de poblacions no seleccionades han estat molt útils per conèixer l'efectivitat de la cirurgia coronària en condicions de pràctica clínica real. Un estudi molt recent realitzat sobre el registre de l'estat de Nova York ha suggerit que el benefici de la revascularització quirúrgica sobre la implementació de stent coronari pot ser aplicable a un grup més nombrós de malalts (en un ventall de risc més ampli) i a més curt termini que el que en un principi suggerien els assaigs clínics (Hannan EL et al, 2005). Aquest estudi suscità un interessant comentari editorial (Gersh BJ et al, 2005) en el qual es comentava que, a efectes d'anàlisi, l'ajust del risc podia tenir un efecte similar a l'assignació aleatòria en magnificar les diferències observades entre els dos grups. En l'estudi esmentat, de manera equivalent al Registre BARI (Detre KM et al, 1999) en el qual els pacients diabètics es tractaven d'acord amb el judici clínic, no s'observaven diferències quan es feia un anàlisi no ajustat, indicant que la tendència a indicar la cirurgia en la majoria de malalts de més risc i la revascularització percutània en la gran majoria de malalts amb malaltia de dos vasos, sense altres criteris de risc, seria doncs apropiada. Aquesta observació subratlla la importància dels estudis observacionals per a l'estudi d'aspectes que un assaig clínic no pot fer sobre la realitat de la pràctica clínica.

La controvèrsia és encara més gran si tenim en compte el ràpid avenç de les tecnologies. És habitual que els assaigs clínics mostrin resultats en períodes que no superen l'any de seguiment, termini probablement insuficient per observar resultats rellevants en determinades pràctiques terapèutiques.

Malauradament però, encara que s'obtinguin resultats a llarg termini (Hlatky MA et al, 2004), aquests queden ràpidament invalidats per la introducció de noves tecnologies que necessiten al mateix temps ésser avaluades. Aquesta forta tensió entre el ràpid avenç tecnològic i la necessitat d'avaluació dels nous tractaments és tanmateix difícil de resoldre; només la interpretació crítica i cautelosa dels estudis publicats, tant dels assaigs clínics com dels estudis observacionals, pot ajudar als professionals mèdics a fer la seva pràctica amb criteris basats en l'evidència.

Els grans canvis que s'han produït en les tècniques de revascularització percutània en els últims anys no han disminuït l'interès en l'estudi de l'efectivitat de la cirurgia coronària doncs, contràriament, això ha augmentat l'exigència dels equips quirúrgics en haver d'afrontar intervencions cada vegada amb major risc associat. Els stents han anat substituint ràpidament l'angioplàstia amb baló (Permanyer-Miralda G et al, 2001) i s'utilitzen ara amb freqüència com a alternativa a la teràpia mèdica o a la revascularització quirúrgica. Els estudis indiquen que els nous dispositius fàrmaco-actius redueixen la necessitat de nova revascularització però no influeixen en la taxa de mortalitat o d'infart de miocardi (Babapulle MN et al, 2004). Hi ha estudis en pràctica clínica real que corroboren i accentuen aquesta troballa dels assaigs clínics: s'ha observat una reducció absoluta de només el 7% del risc de necessitat de nova revascularització (Lemos PA et al, 2004).

També en la tècnica quirúrgica s'han produït canvis importants. Cal destacar la introducció i l'ús cada vegada més generalitzat de la cirurgia sense circulació extracorpòria, que pot suposar una millora en el cost-efectivitat de la cirurgia de revascularització coronària. Encara que els assaigs clínics no han demostrat un benefici clar d'aquesta tècnica sobre resultats "durs" com la mortalitat, la majoria s'han fet sobre poblacions seleccionades, en mostres petites i de risc preoperatori baix (Cheng DC et al, 2005; Nathoe HM et al, 2003). En canvi, en estudis observacionals realitzats en pacients amb un perfil de risc preoperatori més elevat sí que s'observa un avantatge de la cirurgia sense circulació extracorpòria sobre els resultats més rellevants (Wijeysundera DN et al, 2005). D'altra banda falten estudis a més llarg termini que avaluïn amb una metodologia adequada la taxa de noves revascularitzacions.

En aquest marc de canvis fonamentals es fa encara més important l'estudi de l'efectivitat centrat en un context determinat, com una instantània que ens permeti interpretar què està passant, en quin punt ens trobem i quins aspectes es poden millorar. Pel que fa a la cirurgia coronària en particular és plausible suposar que l'evolució dels criteris d'indicació han provocat canvis en el perfil dels pacients operats (s'operen a edats més avançades i amb més comorbiditats) i que, paral·lelament, la introducció de noves tecnologies i l'avaluació dels resultats han induït millores en la tècnica quirúrgica i en el procés assistencial que es tradueixen en millors resultats.

D'altra banda, com més experiència s'adquireix, més complexos són els objectius dels estudis d'efectivitat que, inicialment, es referien tan sols a la xifra de mortalitat en poblacions més o menys àmplies. Posteriorment a això, s'han desenvolupat estudis que volen conèixer aspectes més subtils de la realització de la tècnica quirúrgica en situacions de pràctica clínica real. Per exemple, s'ha analitzat quin és el seu possible efecte sobre la qualitat de vida a curt i llarg termini (Permanyer-Miralda et al, 2001; Herlitz J et al, 1999), o bé, i de manera especialment important, quina és la nostra capacitat de predicció dels resultats clínics (Pons JMV et al, 1997; Hannan EL et al, 1994; Rumsfeld JS et al, 1999).

L'ajust del risc en l'avaluació de resultats

Importància d'ajustar pel risc

La cirurgia coronària segueix essent, malgrat tot, una tècnica no exempta d'un risc important a curt termini o durant l'hospitalització, l'anomenat risc quirúrgic, que cal confrontar amb el benefici esperat. És habitual que els professionals amb experiència, clínics i cirurgians, desenvolupin una noció del risc basada en algunes característiques observades en el malalt individual i jutgin la necessitat de la cirurgia en funció d'aquesta percepció més o menys objectiva. Aquest judici estarà inevitablement influenciat pel caràcter més o menys aversiu al risc del clínic i del cirurgià, i pel grau de participació del malalt en la decisió terapèutica, però la necessitat de predicció del resultat clínic és especialment important.

És evident que el resultat clínic dependrà de diversos factors, a més de l'efectivitat del tractament rebut: de la severitat de la malaltia, de les característiques del malalt que poden fer-lo més o menys susceptible de tenir complicacions, de la qualitat de l'assistència que rebí, l'habilitat, entrenament i criteri tàctic del cirurgià responsable i de moltes altres coses que s'escapen de la nostra capacitat d'observació i que s'agrupen en un calaix de sastre que rep el nom d'atzar.

D'altra banda, les dades sobre resultats clínics s'utilitzen sovint per comparar diferents tractaments o com a indicadors de la qualitat assistencial (Landon B et al, 1996; Iezzoni LI, 1994; Iezzoni LI et al, 1995). En determinats contextos ens pot interessar conèixer la qualitat assistencial dels proveïdors de l'atenció sanitària sigui de manera externa o per l'interès de conèixer els propis resultats. En el cas de la cirurgia coronària, entenem com a proveïdors tant l'equip quirúrgic, com el centre on es realitza, com l'organisme que el gestiona o l'administra.

Tant si estudiem l'efectivitat d'una intervenció, d'una tècnica o d'un tractament concret, com si pretenem comparar el resultat entre diferents

proveïdors, ho voldrem fer bo i exclouent l'efecte de la diferència entre les característiques dels pacients. És per tant necessari conèixer els factors (clínic o no) propis del malalt que s'associen amb el resultat clínic; en el nostre cas concret, amb la mortalitat quirúrgica associada a la cirurgia de revascularització coronària.

L'estratègia d'ajust pel risc mitjançant instruments específics va néixer doncs per a unes finalitats concretes i aquestes son les seves potencials aplicacions pràctiques: 1) la comparació dels resultats ajustats entre proveïdors; 2) l'avaluació dels resultats al llarg del temps; 3) l'ajust del risc en estudis observacionals; 4) l'auto-avaluació crítica dels equips quirúrgics i; 5) la predicció del risc en el malalt individual.

Les escales de risc

Hi ha moltes maneres d'ajustar els resultats per les característiques basals dels pacients. L'assignació aleatòria dels pacients en els diferents grups de tractament a comparar, si és possible, a doble cec, és a dir que ni el pacient ni el metge que el tracta sàpiguen quin és el tractament assignat, és l'única tècnica que teòricament garanteix una distribució de probabilitats d'incidència del resultat clínic idèntica entre els grups. No obstant, en la majoria de situacions reals l'assignació aleatòria no és possible, o només és possible a expenses d'una, de vegades excessiva, falta de validesa externa (els assaigs clínics solen realitzar-se en poblacions seleccionades i poc representatives de la població atesa en condicions reals) o d'una alta ocurrència de *cross-over* (casos que després de l'aleatorització han de ser tractats amb un tractament diferent de l'assignat inicialment).

En el cas de l'avaluació de la qualitat assistencial o de l'observació de la pràctica clínica real l'assignació aleatòria no té sentit. Per això, habitualment s'utilitzen tècniques estadístiques més o menys sofisticades per poder, a partir d'estudis observacionals, controlar la part de variabilitat del resultat deguda a les característiques basals dels pacients. Totes aquestes tècniques permeten l'ajust dels factors de risc observats i registrats, mai d'aquells que no s'hagin observat o que s'hagin passat per alt. En qualsevol cas hem de suposar que les característiques no observades son igualment probables en els grups que pretenem comparar, o almenys cal sospesar la possibilitat de biaixos.

La possibilitat d'aplicar tècniques estadístiques complexes i amb alts requeriments computacionals ha fet que sigui comú l'ús d'anàlisis ajustats (models logístics), en els quals estimem l'efecte d'un factor d'interès sobre el resultat independentment de l'efecte associat a altres característiques basals. L'avantatge principal d'aquest mètode és la possibilitat de controlar per diferents variables o factors de confusió a la vegada obtenint una estimació de l'efecte independent de cada un dels factors sobre el resultat. Es tracta

d'una tècnica robusta i proporciona resultats amb poca probabilitat de biaixos. La possibilitat de biaixos augmenta però quan cal introduir moltes variables en el model en relació al nombre d'esdeveniments.

Alguns autors han utilitzat models basats en el teorema de Bayes (Edwards FH et al, 1988) per la possibilitat que ofereixen de tenir en compte les probabilitats a priori, és a dir la informació disponible a partir d'altres fonts a part de la continguda en les pròpies dades. Altres tècniques sofisticades i amb alts requeriments de càlcul computacional com els sistemes de xarxes neuronals (Lippmann RP et al, 1997) han semblat en alguns moments prometedores però, per ara, no se n'ha generalitzat l'ús.

L'ús dels índexs de propensió (*propensity scores*) està esdevenint popular (D'Agostino RB, 1998). L'índex de propensió és la probabilitat d'un individu de rebre un determinat tractament condicionada a una determinada combinació de possibles factors de confusió. Habitualment es calcula l'índex de propensió mitjançant un model de regressió logística per a la predicció de l'exposició o el tractament d'interès. Com a resultat d'aquesta anàlisi obtenim un valor de probabilitat (l'índex de propensió) que és un resum de tots els possibles factors de confusió. D'aquesta manera podem després utilitzar aquesta puntuació com a "única" variable de confusió de la relació entre l'exposició i el resultat. La possibilitat de tenir en compte moltes covariables a la vegada sense els requeriment de parsimònia que tenen els models logístics tradicionals fa que, si es compleixen determinats requisits, aquesta tècnica sigui especialment interessant en els estudis amb pocs esdeveniments i múltiples factors de confusió (Cepeda MS et al, 2003; Stürmer T et al, 2005).

Finalment, una altra possibilitat és l'ús d'escala de risc estandarditzades en l'avaluació d'aquells procediments per als quals aquestes s'hagin desenvolupat específicament. La literatura sobre models de risc en cirurgia coronària i en cirurgia cardíaca en general és molt abundant. L'elaboració d'aquests models es fa en base a la selecció de determinats factors de risc que es consideren importants segons un judici clínic. Després apareixen o no en el model de risc amb uns pesos més o menys grans segons la seva relació estadística amb la variable de resultat escollida, habitualment la mortalitat hospitalària. Un cop s'ha quantificat la influència de cada factor de risc sobre el resultat, fàcilment es pot fer una predicció del resultat dels pacients coneixent la seva combinació pròpia d'aquests factors de risc. En la majoria de casos els autors solen definir, a partir d'un model logístic, un sistema simple de càlcul que consisteix en sumar una puntuació determinada per la presència de cada factor de risc. D'aquesta manera les puntuacions resulten senzilles d'obtenir sense necessitat de cap aparell de càlcul. El desenvolupament de models de risc de càlcul senzill pot ser molt útil per dues raons: Primera, perquè la construcció de models matemàtics sofisticats com la regressió logística o l'índex de propensió requereix una recollida de dades més exhaustiva i uns

recursos i coneixements que no sempre son accessibles en l'àmbit clínic; i segona, perquè a l'hora de comparar els resultats entre proveïdors o avaluar la qualitat assistencial en un sistema sanitari pot ser molt útil disposar d'una eina comuna que al mateix temps ens serveixi de patró de referència.

En general, aquests models s'obtenen a partir de poblacions àmplies, teòricament representatives d'una zona geogràfica o sistema sanitari i es validen en una altra submostra de la mateixa població o en una mostra d'una població diferent. Tot i que s'ha discutit poc el problema de la població de referència cal recordar que tots els models sense excepció s'han desenvolupat a partir d'unes dades reals que no necessàriament constituïran una referència per a la població en la qual els estem aplicant. Així doncs, quan s'escull un determinat sistema d'ajust del risc és important: 1) conèixer fins a quin punt és vàlida la seva aplicació en el context que es pretén estudiar; 2) reconèixer el perill de circularitat que representa utilitzar la realitat observada com a referència i; 3) acceptar que, en últim terme, l'elecció es basa en un judici de valor respecte de quan i com es consideren acceptables uns resultats determinats.

L'ús d'alguns d'aquests models tant per a la recerca de resultats com per al càlcul del risc en el malalt individual sí que s'ha generalitzat. La seva aplicació generalitzada i l'acceptació de la seva importància va començar a mitjans dels anys 80 quan als EUA es van començar a utilitzar els resultats crus de la cirurgia cardíaca com a indicadors de la qualitat assistencial i els proveïdors, amb tota la raó, van protestar argumentant que no s'havien tingut en compte les diferències en el perfil de risc. El mateix succeïa a més petita escala en les avaluacions internes que es feien dins els propis serveis hospitalaris en el nostre àmbit en aquell temps. A partir de 1989, a l'estat de Nova York, es va començar a recollir, analitzar i disseminar la informació sobre el risc, la mortalitat i les complicacions de la cirurgia coronària (Hannan EL et al, 1990), fet que va significar un important estímul per a la millora de la qualitat en els hospitals de la regió (Hannan EL et al, 1994).

Malgrat la utilitat evident de les escales de predicció del risc per a avaluacions col·lectives en mostres suficientment nombroses, la predicció del risc en el malalt individual està subjecte a elements més imprevisibles i sovint les escales de risc resulten insuficients. La possible utilitat de les escales de risc per a ajudar a establir la indicació d'un procediment estaria en funció de les seves propietats estadístiques de discriminació i de calibració, que haurien de garantir una predicció molt precisa. D'altra banda, cal tenir en compte una limitació inherent a la predicció del risc en el malalt individual (Weintraub WS, 2006) i és que les escales de risc es generen sempre a partir de les dades de pacients en els que la cirurgia ja ha estat indicada, per tant, utilitzar-les com a criteri d'indicació és, si més no, conceptualment enganyós.

Així, en el context nord americà, cal destacar models com el d'Ontario (Tu JV et al, 1995), el model de Parsonnet (Parsonnet V et al, 1989), el de la *Cleveland Clinic Foundation* (Higgins TL et al, 1992), els de la Societat Americana de Cirurgians Toràcics (STS) (Edwards FH et al, 1994) i el recentment desenvolupat específicament per a cirurgia de revascularització coronària a l'estat de Nova York (Hannan EJ et al, 2006;47:661-8).

En l'àmbit europeu el model més utilitzat i que, a més ha estat promogut per les societats científiques europees de cardiologia i de cirurgia cardíaca, és l'EuroSCORE (Nashef SAM et al, 1999) Aquest es va desenvolupar i validar utilitzant les dades sobre el risc i el resultat de les intervencions realitzades l'any 1995 en 128 hospitals de 8 països europeus que participaven en aquesta iniciativa de manera voluntària (Roques F et al, 2000). Es va observar durant aquest període una mortalitat global en les intervencions de revascularització coronària del 3,4% i del 6,8% en el conjunt dels hospitals espanyols que hi van participar. L'EuroSCORE no havia estat validat fins ara per al seu ús en el context de la sanitat pública catalana.

A Catalunya, l'AATRM va desenvolupar un model per a la predicció de la mortalitat de la cirurgia cardíaca (Pons JMV et al, 1997). Es va construir el model a partir de les dades de tots els centres (públics i privats) de Catalunya que realitzaven més de 150 intervencions anuals. Durant l'any 1994 es van fer a Catalunya 1309 intervencions amb una mortalitat quirúrgica del 10,9%. La mortalitat hospitalària per la cirurgia de revascularització coronària aïllada va ser del 8,1%. El model de predicció del risc es va derivar a partir de les dades del 70% de la mostra i es va validar en el 30% restant.

Degut a aquesta gran proliferació de models els investigadors s'han preguntat sovint quin escollir per a la seva aplicació en un context determinat i per determinats objectius, han comparat la validesa dels més utilitzats (Iezzoni LI et al, 1998) i han desenvolupat models locals. Malgrat tot, encara hi ha controvèrsia respecte quina és l'aproximació òptima a l'hora d'ajustar pel risc (Peterson ED et al, 2000). Per exemple, en quines situacions és millor utilitzar un instrument d'àmplia difusió com l'EuroSCORE o un model d'àmbit local com seria en el nostre cas el model de l'AATRM (Pons JMV et al, 2000; Pitkänen O et al, 2000), o quan és necessari actualitzar les escales disponibles (Ivanov J et al, 1999; Krumholz HM, 1999).

Lògicament, quan s'han comparat les característiques de validesa (discriminació i calibració) entre models externs i models locals de recent desenvolupament per a la seva aplicació en el context local (Pons JMV et al, 2000; Pitkänen O et al, 2000; Ivanov J et al, 1999), els segons s'ajusten sempre millor als resultats observats.

L'avantatge més evident de l'ús d'un model d'ampli abast és la possibilitat de fer comparacions internacionals però, en canvi poden no tenir en compte les

peculiaritats pròpies de la població en un context local i diluir les diferències que hi poden haver entre diferents sistemes sanitaris. En canvi, els models derivats a partir d'una mostra representativa d'un sistema sanitari concret poden ser més útils per a avaluacions locals i per a interpretar canvis en el temps.

L'estudi CIRCORCA i la influència del tipus de gestió hospitalària

A Catalunya la cirurgia coronària es porta a terme en centres de gestió pública i de gestió privada. En alguns d'aquests últims, però, a més de malalts procedents d'assegurances privades i de malalts estrictament privats, s'opera una proporció considerable de malalts atesos a la sanitat pública i remesos als esmentats centres tan sols per a la pràctica de l'operació quirúrgica.

En l'estudi CIRCORCA es van incloure de manera prospectiva tots els pacients sotmesos a una primera intervenció d'empelt coronari, sense altres operacions associades, amb finançament públic o privat, en els 8 hospitals de gestió pública i privada que feien en aquell moment més de 100 intervencions anuals, entre novembre del 1996 i juny del 1997. A més de recollir la informació sobre el risc preoperatori i el resultat clínic a curt termini, es va valorar el perfil de qualitat de vida relacionada amb la salut (QVRS) preoperatori i als sis mesos i a l'any de la intervenció, utilitzant dos instruments: un qüestionari genèric (SF-36) i un qüestionari específic per valorar la capacitat funcional en malalts cardíopates (Duke Activity Status Index).

Vegeu a l'apartat Annex I els dos articles publicats sobre l'estudi CIRCORCA.

L'estudi CIRCORCA ens va permetre conèixer que a Catalunya la mortalitat quirúrgica d'una primera derivació aortocoronària era alta, bàsicament atribuïble al risc quirúrgic preoperatori moderat o alt en una proporció important de la població de l'estudi, però significativament més elevada entre els malalts procedents de la sanitat pública que en els procedents de la sanitat privada (8,2% vs 0,7%). Els predictors independents de la mortalitat hospitalària, a més del tipus de finançament públic, eren l'edat, la classe funcional III o IV en el moment de la intervenció, el caràcter urgent o emergent de la cirurgia, la presència d'insuficiència ventricular esquerra, la necessitat de baló intraaòrtic i la presència de malaltia vascular perifèrica (Permanyer-Miraldà G et al, 1998).

Quan es va fer el seguiment a l'any (Permanyer-Miraldà G et al, 2001) es va comprovar que la diferència de mortalitat entre els dos subgrups segons el tipus de finançament tendia a reduir-se, essent, però, encara significativa (11% dels pacients amb finançament públic i 5% dels pacients amb finançament

privat). A més, en l'anàlisi multivariada, el fet de procedir de la sanitat pública fou també predictor independent de la mortalitat tardana juntament amb la mala qualitat de vida preoperatoria, la presència de comorbiditat, l'edat i un indicador de revascularització incompleta.

De forma interessant el finançament públic, el gènere femení, la presència de comorbiditat i la pitjor qualitat de vida basal eren predictors de pitjors nivells de qualitat de vida a l'any de la intervenció. S'observava que, malgrat que la millora es produïa en tots els grups, aquests ja partien de pitjors nivells de qualitat de vida basal.

A més, l'estudi va comprovar una mortalitat significativament més alta en els centres públics que en els privats. Atès que dos dels centres de gestió privada inclosos a l'estudi (Centre Cardiovascular Sant Jordi i Hospital General de Catalunya) operaven una proporció de 30% i 80% de malalts, respectivament, procedents de la sanitat pública, es va poder analitzar per separat la mortalitat dels malalts operats en aquests centres segons el seu tipus de finançament de l'atenció sanitària. Es va comprovar que, mentre que en els malalts procedents de la sanitat privada la mortalitat era molt baixa (0,5%), en els procedents de la sanitat pública era clarament superior (6,8%) però sense arribar al valor observat en els malalts de la sanitat pública operats en centres públics (8.4%). La diferència entre aquests dos valors no era significativa, però la mostra de malalts públics operats en centres privats era insuficient.

El conjunt de les troballes de l'estudi CIRCORCA no permetia afirmar si la diferència de mortalitat hospitalària observada entre els centres de gestió pública i els de gestió privada era només atribuïble a les diferències basals entre el risc preoperatori d'ambdós subgrups, o bé si hi influïa algun factor relatiu a un diferent procés assistencial en els dos tipus de centre. Aquest fet ens va portar a dissenyar l'estudi ARCA (Avaluació dels Resultats de la Cirurgia Coronària a Catalunya), en el qual obtindríem una mostra més àmplia d'intervencions amb finançament públic realitzades en centres de gestió privada i en el que utilitzaríem instruments específics per a l'ajust del risc preoperatori.

És interessant ressaltar, per altra banda, que, de manera pràcticament contemporània, en els treballs de validació de l'instrument EuroSCORE (Nashef SAM et al, 1999), dissenyat per a avaluar el risc de la cirurgia cardíaca específicament en països europeus, es va comprovar que la mortalitat global crua i ajustada) de la cirurgia cardíaca en la mostra espanyola, bàsicament procedent de la sanitat pública, era la més alta dels sis països europeus participants. La taxa de mortalitat era, només a la mostra Espanyola, clarament superior a la que predeïa l'instrument. La mostra espanyola posseïa també l'índex més baix d'ús d'empelt d'artèria mamària interna, que pot ser considerat un criteri de qualitat tècnica.

OBJECTIUS

- 1) Avaluar la mortalitat hospitalària de la cirurgia de revascularització coronària, ajustada al risc preoperatori, en malalts de la sanitat pública operats en centres de gestió pública i privada de Catalunya.
- 2) Revalidar i comparar el funcionament, en la mateixa població, de dos models d'avaluació del risc de mortalitat després de la cirurgia de revascularització coronària:
 - Una escala desenvolupada ja fa una dècada en el mateix àmbit local on es fa ara la seva aplicació: L'escala de l'AATRM.
 - Una escala desenvolupada en l'àmbit europeu, un àmbit més ampli que el de la seva aplicació actual: l'EuroSCORE.
- 3) Com a objectiu secundari, comparar, en la mateixa població, el resultat de la cirurgia sense CEC amb la cirurgia amb CEC en pacients de risc alt i baix operats en la pràctica clínica habitual en cinc hospitals públics i privats de Catalunya.

MÈTODES I RESULTATS

Els tres estudis s'han realitzat a partir de les mateixes dades i per tant, el disseny, la població d'estudi i l'estratègia de recollida de dades son comuns a tots ells.

Disseny i població d'estudi

Es van reclutar de forma prospectiva i consecutiva tots els malalts programats per a una primera intervenció d'empelt coronari aïllat, entre novembre de 2001 i novembre de 2003, en cinc hospitals de Catalunya (tres grans hospitals públics i els dos únics hospitals de gestió privada que operen malalts procedents de la sanitat pública).

Estratègia de recollida de dades i variables de l'estudi

Vegeu l'apartat Annex 2 el formulari de recollida de dades i les definicions de les variables d'estudi i a l'apartat Annex 3 la informació detallada sobre les dues escales de risc utilitzades (l'EuroSCORE i l'escala de l'AATRM)

La recollida de dades es va realitzar per un investigador pertanyent a l'equip de cardiologia o de cirurgia cardíaca de cada un dels centres participants. Després de dues reunions de l'equip investigador i de realitzar una prova pilot que consistí en la inclusió i registre de cinc casos en cada centre participant, es va elaborar la versió definitiva del protocol d'estudi, el qüestionari de recollida de dades i el manual de definicions de les variables de l'estudi.

La recollida de dades es feia en dos moments en el temps:

En el moment de la indicació quirúrgica es recollien les característiques socio-demogràfiques (sexe, edat, situació laboral, nivell d'estudis), factors de risc cardiovascular (hipertensió arterial, hipercolesterolèmia, diabetis, tabaquisme), antecedents de malaltia cardiovascular (antecedents clínics de cardiopatia isquèmica, revascularització, insuficiència cardíaca o accident vascular cerebral), les característiques de la malaltia coronària que motivava la indicació (nombre de vasos afectats, afectació de la descendent anterior proximal o del tronc comú) i les característiques basals de risc segons les definicions de l'EuroSCORE i l'escala de l'AATRM (creatinina elevada, disfunció neurològica, malaltia hepàtica, malaltia pulmonar crònica, aneurisma ventricular esquerre, claudicació intermitent, oclusió o estenosi de >50% d'una arteria caròtida, intervenció prèvia o programada de l'aorta abdominal, caròtides o extremitats, classificació del grau funcional segons la Canadian

Cardiovascular Society, hipertensió pulmonar sistòlica, disfunció ventricular esquerra, infart de miocardi previ i data de l'últim infart si n'hi havia.

Durant l'ingrés per la intervenció es recollien els esdeveniments ocorreguts durant el temps d'espera (ingrés per infart o angina inestable, nombre d'ingressos o atencions a urgències per malaltia coronària, mort per causa cardiovascular o no, refús o demora de la intervenció i motiu de la demora), les característiques prequirúrgiques de risc segons les definicions de l'EuroSCORE i l'escala de l'AATRM (a més de les recollides en el moment de la indicació, el caràcter urgent o emergent de la intervenció, la presència de crisi d'angina de repòs en les 72 hores prèvies a la intervenció, xoc cardiogènic, necessitat de nitrats intravenosos fins a l'entrada a quiròfan, taquicàrdia o fibril·lació ventricular, massatge cardíac preoperatori, intubació abans de l'entrada a quiròfan, suport ionotròpic intravenós preoperatori, necessitat de baló de contrapulsació intraaòrtica, insuficiència renal aguda en el preoperatori i ruptura septal postinfart), les variables relacionades amb la tècnica quirúrgica realitzada (nombre d'empelts arterials de mamària o radial o de safena, cirurgia sense circulació extracorpòria) i finalment els esdeveniments durant el postoperatori immediat o 30 dies després de la intervenció (sagnament, infart, reintervenció, complicacions infeccioses nosocomials, respiratòries o de la ferida esternal, altres complicacions greus cardiovasculars o no cardiovasculars, o mort).

Finalment, per determinar la mortalitat a tres mesos i confirmar la mortalitat a 30 dies de la intervenció es va creuar la base de dades amb el Registre de Mortalitat del Servei Català de la Salut.

Per assegurar la qualitat de les dades es van realitzar reunions de treball aproximadament cada quatre mesos durant el curs de l'estudi per tal de solucionar tots els problemes relacionats amb el reclutament de malalts i el registre de les variables d'estudi. Totes les definicions de les variables d'estudi es van consensuar amb els investigadors de cada centre encarregats de la recollida de dades i es van incloure en un manual de definicions. A més d'analitzar de manera periòdica les dades per tal de detectar possibles errors es va fer, en dues ocasions al llarg del període d'inclusió, una selecció aleatòria del 10% dels casos inclosos en cada centre per fer-ne la revisió de les històries clíniques per personal extern al centre. Per assegurar que la recollida de dades era exhaustiva es van comparar els casos inclosos a l'estudi amb els casos inclosos en el CMBD durant el mateix període en els que constava un procediment de cirurgia de derivació aortocoronària (codi 36.1x de l'índex de Codis Diagnòstics ICD-9).

ARTICLES

	Referència	Objectius
Article 1	Ribera A, Ferreira-González I, Cascant P, Pons JMV, Permanyer-Miralda G por el grupo de investigadores del estudio ARCA. Evaluación de la mortalidad hospitalaria ajustada al riesgo de la cirugía coronaria en la sanidad pública catalana. Influencia del tipo de gestión del centro (estudio ARCA). Rev Esp Cardiol 2006;59:431-40	Avaluar la mortalitat hospitalària de la cirurgia de revascularització coronària, ajustada al risc preoperatori, en malalts de la sanitat pública operats en centres de gestió pública i privada.
Article 2	Ribera A, Ferreira-González I, Cascant P, Pons JMV, Permanyer-Miralda G for the ARCA study group investigators. The EuroSCORE and a local model consistently predicted coronary surgery mortality and showed complementary properties. J Clin Epidemiol (en premsa)	Revalidar i comparar el funcionament de dos models d'avaluació del risc de mortalitat després de la cirurgia de revascularització coronària: 1) Una escala desenvolupada ja fa una dècada en el mateix àmbit local on es fa ara la seva aplicació: L'escala de l'AATRM. 2) Una escala desenvolupada en l'àmbit europeu, un àmbit més ampli que el de la seva aplicació actual: l'EuroSCORE.
Article 3	Ferreira-González I, Ribera A, Cascant P, Permanyer-Miralda G for the ARCA study group. Outcomes in off-pump versus on-pump coronary artery bypass grafting stratified by operative risk profile: an assessment using propensity score. Eur Heart J 2006;27:2473-80	Comparar el resultat de la cirurgia sense CEC amb la cirurgia amb CEC en pacients de risc alt i baix operats en la pràctica clínica habitual en cinc hospitals públics i privats de Catalunya

ARTICLE I

Ribera A, Ferreira-González I, Cascant P, Pons JMV, Permanyer-Miralda G por el grupo de investigadores del estudio ARCA. Evaluación de la mortalidad hospitalaria ajustada al riesgo de la cirugía coronaria en la sanidad pública catalana. Influencia del tipo de gestión del centro (estudio ARCA). Rev Esp Cardiol 2006;59:431-40

Ribera A, Ferreira-Gonzalez I, Cascant P, Pons JM, Permanyer-Miralda G.

Evaluación de la mortalidad hospitalaria ajustada al riesgo de la cirugía coronaria en la sanidad pública catalana (estudio ARCA)

Revista Española de Cardiología. 2006 May;59(5):431-40.

ARTICLE 2

Ribera A, Ferreira-González I, Cascant P, Pons JMV, Permanyer-Miralda G for the ARCA study group investigators. The EuroSCORE and a local model consistently predicted coronary surgery mortality and showed complementary properties. *J Clin Epidemiol* (en premsa)

Date: Oct 02, 2006
To: "Gaietà Permanyer-Miralda" gpermany@vhebron.net
cc: kim.luyten@hag.unimaas.nl
From: kim.luyten@hag.unimaas.nl

Subject: JCE-06-177R2: Final Decision
Ms. Ref. No.: JCE-06-177R2
Title: The EuroSCORE and a local model consistently predicted coronary surgery mortality and showed complementary properties
Journal of Clinical Epidemiology

Dear Dr. Permanyer-Miralda,

The editors are pleased to inform you that your manuscript "The EuroSCORE and a local model consistently predicted coronary surgery mortality and showed complementary properties" has been accepted for publication in Journal of Clinical Epidemiology.

The pertinent material will be forwarded to the publisher who should send you galley proofs, for your approval and expedient return (to the publisher), within the next few months. The copyright forms requested by the publisher will be included with the galley proofs.

Sincerely yours,

J. André Knottnerus, Editor

Kim Luyten
Editorial Assistant

Journal of Clinical Epidemiology
Maastricht Editorial Office
TEL: +31 43 3882319
FAX: +31 43 3671458
E-mail: kim.luyten@hag.unimaas.nl

The EuroSCORE and a local model consistently predicted coronary surgery mortality and showed complementary properties

Authors: Aida Ribera (1, 2), Ignacio Ferreira-Gonzalez (1, 3), Purificació Cascant (1), Joan MV Pons (4), Gaietà Permanyer-Miralda (1), for the ARCA Study group investigators.

Authors affiliations: (1) Epidemiology Unit. Cardiology Department. University Hospital Vall d'Hebron. Barcelona, Spain. (2) Department of Health and Experimental Sciences. Universitat Pompeu Fabra. Barcelona, Spain. (3) Department of Medicine. Universitat Autònoma de Barcelona, Spain. (4) Catalan Agency for Health Technology Assessment. Barcelona. Spain

Complete author information:

Aida Ribera

Epidemiology Unit. Cardiology Department. University Hospital Vall d'Hebron

Psg. Vall d'Hebron, 119-129 08035 Barcelona Spain

Tel. 00-44-932746177 / Fax. 00-44-932746063 / aribera@vhebron.net

Ignacio Ferreira-Gonzalez

Epidemiology Unit. Cardiology Department. University Hospital Vall d'Hebron

Psg. Vall d'Hebron, 119-129 08035 Barcelona Spain

Tel. 00-44-932746177 / Fax. 00-44-932746063 / iferreir@vhebron.net

Purificació Cascant

Epidemiology Unit. Cardiology Department. University Hospital Vall d'Hebron

Psg. Vall d'Hebron, 119-129 08035 Barcelona Spain

Tel. 00-44-932746177 / Fax. 00-44-932746063 / pcascant@vhebron.net

Joan MV Pons

Catalan Agency for Health Technology Assessment

Esteve Terradas, 30. Recinte Parc Sanitari Pere virgili. Edifici Mestral, 1a planta. 08023 Barcelona. Spain

Tel. 00-44-932594200 / Fax 932594201 / jpons@aatrm.catsalut.net

ARCA Study group investigators:

Marcos Murtra (Department of Cardiac Surgery. University Hospital Vall d'Hebron. Barcelona, Spain)

Carlos Sureda, Miguel Ángel Castro and Xavier Ruyra (Department of Cardiac Surgery. Hospital Germans Trias I Pujol. Barcelona, Spain)

Bernat Romero, Francisco Callejo and Alejandro Arís (Hospital de la Santa Creu i de Sant Pau. Barcelona, Spain)

Carmen Martínez-Useros and Jaume Mulet (Hospital General de Catalunya. Barcelona, Spain)

Mireia Sanz and Oriol Solé (Centre Cardiovascular Sant Jordi. Barcelona, Spain)

Word count: 3,547

Abstract

Objective: To revalidate a local model for prediction of in-hospital mortality after coronary surgery several years after its introduction and the EuroSCORE in a specific area within its original scope. To assess the specific advantages of one type of instrument over the other in a definite context.

Study design and Setting: Data from consecutive patients undergoing a first isolated coronary artery bypass between November 2001 and November 2003 in five hospitals in Catalonia were prospectively collected.

Results: 1,605 patients were included. Area under the ROC curves were around 0.75 for both models. Calibration was low for both models and the local model significantly overestimated risk. The ordering of operating centres by performance was identical with each strategy but the centers labelled as outliers differed.

Conclusion: 1) Evaluation of performance of individual hospitals was consistent using both systems and almost identical when they were internally recalibrated. 2) The impact of the benchmark population characteristics on model performance may be greater than that of risk factors considered for score calculation. 3) Promoting the use of a widely used instrument as the EuroSCORE might be sufficient for most evaluations. Local scales can be useful to highlight locally relevant features and temporal trends.

Keywords: coronary artery surgery; surgical risk; mortality; EuroSCORE; risk assessment; validity

Running title: Comparison of a locally derived model and the EuroSCORE for risk assessment in coronary artery surgery.

Abstract word count: 199

Introduction

Risk scores are an important tool for the assessment of outcomes of coronary artery bypass graft (CABG) and other clinical procedures. Several risk-adjustment methods have been developed in the last twenty years to predict in-hospital mortality following cardiac surgery. Although much valuable information has been gathered using several instruments, an agreement as to the optimal instrument to be used has not been achieved. It is still controversial whether a universal or widely applicable instrument should be developed, or when locally derived ones are more appropriate [1]. Although the different models might differ on the methods for developing the algorithms, in essence all score systems are based on the same patient preoperative data, such as age, comorbidity and severity of disease. Therefore, acknowledging the differences between the context in which the scores are developed and where it is subsequently applied is crucial to ensure the validity of its application in real practice. For instance, it has been observed that predictive models developed in the era of conventional angioplasty cannot be applied directly to current interventional practice with stent implantation [2]. In a similar way, changes in treatment of acute coronary syndromes, as well as changes in surgical techniques may have modified the practice of CABG.

The index which has been promoted by the European Societies of Cardiology and Cardiac Surgery for its use in Europe is the EuroSCORE, as it was developed and validated in a wide population from eight European countries [3]. The advantage of a method which is intended for general application in a broad context is the possibility of European-wide comparisons but, on the other hand, it can fail to take into account the particularities of the population in a local context and dilute the differences that may exist between health care systems. Although the EuroSCORE has been validated in individual European countries [4] evaluating its performance in other patient populations is still needed.

At the same time, locally derived models are currently being developed with the aim of getting more precise predictions in a particular health system [5]. Such methods might be more useful for local evaluations and also for better interpretation of changes over time in a particular context. The differences between broad and locally derived models would thus rely, not in the risk factors they consider, but in the different reference populations used as a benchmark for mortality assessment.

The aims of the present study were the following: 1. To revalidate both the Catalan Agency for Health Technology Assessment (CAHTA) risk model (a locally derived model developed for its use in Catalonia) several years after its introduction and the EuroSCORE in a specific area within its original scope. 2. To assess the specific advantages of one type of instrument over the other type in a definite context.

Methods

Data collection

Data from all consecutive patients from the public health system in whom a first coronary artery bypass surgery without associated procedures was indicated were prospectively collected from November 2001 through November 2003 in three major teaching public hospitals and in the two privately managed institutions that admit patients from the public system in Catalonia. Patients whose health care was funded by private health insurance companies or who paid for it out-of-pocket were explicitly excluded.

In all centres a member of the cardiological or surgical teams reviewed the operation schedules and fulfilled the questionnaire by means of patient interview, review of clinical records and discharge reports. The research protocol was approved by the local ethics committee and informed consent was obtained from all subjects. The quality of the data collection was checked by a review of randomly selected clinical records (10% of the recruited sample) by the research team, and by a head-to-head comparison of the list of included patients with the administrative databases.

Risk assessment

The risk of hospital mortality was assessed using two predictive instruments that have been validated in several settings using logistic regression methods. The EuroSCORE was developed and validated [3, 4] in eight European countries to predict the risk of surgical mortality. More recently the logistic equation that had been used to develop the risk adjustment model has been reported, so that surgical mortality can be more accurately predicted using this equation [6]. To increase the objectiveness of the variable 'unstable angina', data collectors were asked to include an additional variable reporting angina at rest within the 72 hours previous to operation. Angina was considered 'unstable' only when these criteria were fulfilled. For all other variables definitions were as used according to the descriptions of both instruments [3, 5].

The CAHTA model was developed in Catalonia in 1994 [5] for open-heart procedures including valvular, coronary and mixed procedures.

Statistical analysis

Prevalences of all risk factors defined in both risk adjustment methods and adjusted odds ratios for inhospital mortality in our study population were calculated. Data from the development sample of the CAHTA risk model and its logistic regression equation were obtained directly from the authors [5] and the subgroup of patients sharing the same inclusion criteria as the study population were selected for comparison of risk factor prevalences between the two samples.

To estimate the predicted probabilities of inhospital death, the logistic regression equations of the two original models [5, 6] were calculated for each patient in the study population.

Validity of both models was assessed by evaluating statistical performance and evaluative usefulness (or the model's influence to risk-adjusted outcome ratings for individual hospitals).

Statistical performance was evaluated by means of model discrimination and calibration. Discrimination indicates the model's capacity to differentiate patients who will develop a given outcome from patients who will not. Calibration refers to the degree of concordance between model's risk prediction and actual incidence of outcome in the whole risk range.

Discrimination was assessed with the area under the receiver-operating characteristics (ROC) curve. Comparison of ROC curve areas for the two models was performed according to the method of DeLong and cols [7].

Calibration was assessed by the Hosmer-Lemeshow goodness-of-fit statistic: the data are grouped in quantiles of the model estimated probabilities and observed versus expected deaths in each quantile are compared with a χ^2 statistic. As differences between observed and expected number of deaths increase, the χ^2 value increases and the p value decreases, thus, a lower p value indicates worse model calibration. To assess whether the relationship was overestimation or underestimation we inspected the plots of observed versus predicted mortality in each quantile [8], we estimated the parameters of this relationship with weighted linear regression and tested whether the slope was equal to 1 (which is the expected slope when the relationship is optimal) with a Wald test.

Significant overprediction or underprediction can occur when event rates in the study sample are lower or higher than in the development sample. In this case recalibrating the model can be considered and different methods for recalibration are available [8]. We performed 'recalibration-in-the-large' [9] by adjusting a new logistic regression to fit an intercept term (α) and a multiplier term (β) to the original risk score (RS), so that the revised predicted probability (P) can be predicted as follows:

$$P = \exp(\alpha + \beta RS) / [1 + \exp(\alpha + \beta RS)]$$

where P is the revised predicted probability of in-hospital mortality, RS is the original risk score and α and β are estimated when the model is applied in the current population. This approach simply implies adjusting the overall mean and slope and does not imply re-estimation of the coefficients for each individual risk factor.

To assess the evaluative usefulness (or the model's influence to risk-adjusted outcome ratings for individual hospitals), the mortality rate in each hospital was adjusted for risk using the ratio between observed and expected deaths. Outlier hospitals were identified as those in which observed/expected 95% confidence bounds excluded unity. In addition hospitals rank-ordering by risk-adjusted mortality rates were qualitatively compared across the two models.

Results

Out of 1,640 patients from the public health system in whom surgery had been programmed, 1,605 were operated on (412 cases in hospital 1, 410 in hospital 2, 445 in hospital 3, 138 in hospital 4 and 200 in hospital 5). In table 1 and 2 both scales used for risk assessment are described. In table 1, prevalences and risk factor adjusted odds ratios for inhospital mortality in the derivation study [5] and in our study population (which serves as a revalidation sample) are presented to illustrate changes in the risk profile of patients operated in the two study periods.

(Tables 1 and 2)

Overall observed mortality was 4.8%. Remarkably, all the OR in the development sample fell within the confidence limits of the OR in our study population. Only the OR for unstable angina was higher in our study population than in the derivation sample for the EuroSCORE (table 2).

Figure 1 shows the ROC curves for the CAHTA risk model and the EuroSCORE. The discriminative value was fairly good (area under the ROC curves around 0.75) for both models [10] and the differences between the two areas were not statistically significant.

(Figure 1)

Calibration was low for both models and the p value for the Hosmer-Lemeshow χ^2 statistic was below 0.05 for the CAHTA model. Figure 2 depicts the plot of observed deaths versus expected in each decile of estimated probabilities for inhospital death by the CAHTA risk model and the EuroSCORE. The CAHTA model overestimated the number of deaths in eight of the nine decile groups and the slope was significantly higher than 1 indicating significant overestimation. For the EuroSCORE the observed number of deaths was significantly higher than expected in the lower risk decile group and for the CAHTA risk the observed number of deaths was significantly lower than expected in the higher risk decile group. When both models were internally recalibrated perfect fit was obtained.

(Figure 2)

In figure 3, adjusted mortality (the ratio of observed vs expected mortality using both models) for each participating hospital are represented. The ranking of hospitals based on their adjusted mortality was the same when using both systems for risk adjustment (Figure 3, A and B) except for hospitals 2 and 3 which had virtually identical adjusted mortality.

(Figure 3)

Using the CAHTA model as the risk adjustment tool (Figure 3A), all hospitals except one had an adjusted mortality with the upper limit of the 95% confidence interval well below 1, indicating that the observed mortality was significantly lower than expected. Using the EuroSCORE as the risk adjustment tool, all hospitals except one had 95% confidence intervals including 1. Hospital 1 had lower 95% confidence interval limit over 1, indicating that the observed mortality was significantly higher than expected.

When the models were internally recalibrated (figure 3, C and D) the consistency between both models becomes even more apparent. The

estimated intercept term (α) and multiplier term (β) for recalibration of the CAHTA risk score were: $\alpha = -4$ and $\beta = 0,06$; and for the EuroSCORE: $\alpha = -4,7$ and $\beta = 0,12$. Using this approach Hospital I had a mortality rate significantly higher than expected using the recalibrated CAHTA risk score. Using the recalibrated EuroSCORE all hospitals had 95% confidence intervals including I.

Discussion

In the context of a study designed to evaluate the results of coronary artery bypass surgery in a whole health care region we used two different tools for risk assessment: the EuroSCORE and a locally derived score, the CAHTA model.

Both models needed revalidation for different reasons: the EuroSCORE was applied in a specific area within its original scope that may have different features from the average EuroSCORE population, and the CAHTA was applied several years after its introduction. So that, the aim of the present study was, first, to revalidate both instruments and, second, to compare their performance in a definite context. We approached both by assessing and comparing statistical performance and evaluative usefulness of both systems.

Statistical performance

While discrimination abilities of both models were good, with areas under the ROC curves similar to those obtained in the original development and validation studies, calibration was poor, especially for the CAHTA model, which showed significant overestimation. The calibration of the EuroSCORE was better but not optimal as it tended to underestimate mortality in the high risk patients.

Recalibration (re-estimation of the coefficients for each risk factor) of the local scale would be a good strategy in this situation [11], but the sample size of the present study was not enough to ensure stable models when splitting the sample into development and validation subsamples. Moreover, as shown in table I, the only factors in the CAHTA model showing a significant effect on in-hospital mortality were age, functional class IV and elevated creatinine; thus, improving the index would probably imply not only recalibration, but also considering other variables such as comorbidities (chronic pulmonary disease and extracardiac arteriopathy) and more objective measures of clinical stability.

When a simpler strategy of recalibration (recalibration-in-the-large) was used instead, perfect fit was obtained between observed and expected mortality with both recalibrated models. This strategy, which implies calculating a new risk score by adding a constant and a multiplicative factor to the old score, should be considered when the lack of calibration is important.

Evaluative usefulness

Comparison of adjusted mortality between providers or between different time periods represents one of the applications of risk assessment in real practice. Even allowing for risk overestimation of the CAHTA model, both methods for risk adjustment provided the same ranking of the participating centres when ordering them by adjusted mortality values. Nevertheless, as it has also been demonstrated in other clinical settings [2, 12], differences in model calibration markedly influenced which hospitals were labelled as outliers. This finding illustrates that old models, or models developed outside the context where they were generated, work reasonably well as long as they are not used to identify outliers [13]. Otherwise, recalibration of the scales should be considered.

After internal recalibration of the two models, the adjusted mortality estimates became almost identical (figure 3) and, for the CAHTA score, substantially different from values before recalibration. This suggests that the reference population used as a benchmark has a great impact on model performance, this influence being even greater than that of the risk factors considered for the score calculation. Thus, interpretation of this influence deserves further attention as it might help choosing the best model for risk assessment.

Benchmark appropriateness

The validity for our setting of the reference population from which the two instruments were derived is a fundamental issue. As for the CAHTA model, it was built ten years ago, at a time when there are some reasons to suppose that surgical results were poorer than the present ones in the study area [14, 15], so it could be anticipated that the instrument would overestimate risk. With regard to the EuroSCORE, it might be wondered to what extent an instrument derived from a heterogeneous pool of European populations several years ago applies to our current results. In fact, in a previous study [4], surgical mortality in Spain adjusted by the EuroSCORE was higher than expected, in marked contrast with other European countries. One possible explanation could be that the instrument was not suitable for the Spanish population. However, we think that these findings are best explained by truly poorer results in the included Spanish hospitals at that time [16]. In fact, in the present study there is a remarkable agreement between both instruments even allowing for overestimation by the CAHTA model: the ranking of the five centers for adjusted mortalities is also the same with both instruments.

Inclusion of one hospital with a high weight (26 % of the overall sample) and an observed mortality significantly higher than expected by the logistic EuroSCORE might have increased its tendency to underestimation. This fact can be interpreted as a limitation for the validation process but at the same time it underscores the usefulness of the instrument for evaluative purposes in detecting outliers.

As we have seen, for evaluative purposes it is important to recognize the differences in mortality between the benchmark and the study population. Equivalence can only be attained by refitting new models or by recalibrating the existing models [11]. Alternatively there is also the possibility of performing 'recalibration-in-the-large', which uses an internal benchmark in a way that outliers are identified in relation to their peers and not to an external and possibly incomparable benchmark.

Relevant risk factors

Once accepting that the selection of a benchmark is important for further interpretation of results, it seems reasonable to choose a model which is clinically meaningful and appropriate for the setting and purpose of the study.

It is well known that instrument scales cannot take in all potentially relevant risk factors, some of which cannot easily be quantified but may be particularly prevalent in a given population, such as physical frailty, or those important for a given individual but not frequent in most populations (e.g. immunosuppressant therapy or hepatic dysfunction). The relevance of this type of influence is illustrated by a study where the subjective assessment of risk by surgeons was more precise than by using the risk instrument in the upper and lower ends of the scale, the converse being true [17]. Should any of these unrecorded variables have been common in our patient population, a reason for their consideration in the risk analysis would have arisen, with the contrasting needs of comprehensiveness and parsimony.

Some of the differences between the two approaches might be due to the different definition of risk factors in the derivation studies of both scales. E.g. both scales in the present study have different indicators of unstable angina: in the CAHTA score it is not clear how and when the definition of functional class is applied and in the EuroSCORE, although the definition of unstable angina is implied as the need for endovenous nitrates, this might be influenced by local practices in each hospital, while how long after the last anginal episode should the variable "unstable angina" apply is not clear. The influence of unstable angina is of great impact in our setting as it was a strong predictor of inhospital mortality (the OR for unstable angina was 2.7, and the OR for functional class IV was 2.6). On the other hand, all the ORs which determined the score weights for the CAHTA score were within the 95% confidence intervals of the ORs in our study population, while this is not true for unstable angina in the EuroSCORE, which apparently has a greater impact on mortality in our context (OR=2.7) than it had in the EuroSCORE development sample (OR=1.5). This might illustrate one of the advantages of using a local model to reflect more closely the true weight of a variable differently weighted in other models.

Appropriate benchmarking and precise and unambiguous definitions of predictor variables are essential to interpret changes over time in the weight of variables as an improvement of surgical practices. Moreover, a risk assessment method implemented in a nationwide well established database

which allows for periodic recalibration or refitting might be the best approach for continuous quality improvement activities [18, 19].

In summary, the advantage of using the two approaches for risk assessment, a local scale and a Europeanwide scale, was in our experience more tactical than conceptual. By using the two systems we ensured that all relevant risk factors were collected and permitted both, comparison with past results in the same health system and international comparisons within the European context. Moreover, using the two instruments in real practice is efficient as most features are common to both and the effort for collecting them is affordable once the system for registering presurgical risk information is implemented.

Limitations

Unlike in our study, the original studies for derivation and validation of both models included data on reoperations and other cardiac surgery procedures as well as coronary artery surgery alone. This makes comparison of the populations more difficult to interpret as the risk factors prevalences might be different due to the differing selection criteria. However, the adjusted effect of the same risk factors on inhospital mortality should be comparable regardless of their prevalence, and important differences on the odds ratios, as we have discussed, would indicate differences of the relative importance of the risk factors for predictions.

Conclusions

1) When used in a representative sample of a whole health care region the EuroSCORE had an acceptable statistical performance and the CAHTA score (a locally derived model) showed a significant overestimation of risk. Nevertheless, evaluation of performance of individual hospitals was consistent using both systems and almost identical when both models were internally recalibrated.

2) The impact of the benchmark population characteristics on model performance may be greater than that of risk factors considered for score calculation.

3) Promoting the use of a widely used instrument as the EuroSCORE might be sufficient for most evaluations. But developing local scales for periodical evaluation of results in a local context (such as a whole health care region within a country) can be useful to highlight the influence of locally relevant features and to understand temporal trends and changes in practice.

Acknowledgements

This study was supported by funds from the Catalan Agency for Health Technology Assessment (CAHTA 061/22/2000) and partially for the Instituto de Salud Carlos III (network of excellence Red IRYSS G03/202).

References

1. Pitkänen O, Niskanen M, Rehnberg S, Hippeläinen M, Hynynen M. Intra-institutional prediction of outcome after cardiac surgery: comparison between a locally derived model and the EuroSCORE. *Eur J Cardio-thorac Surg* 2000;18:703-710
2. Kizer JR, Beerlin JA, Laskey WK, Sanford Schwartz J, Sauer WH, Krone RJ, Kimmel SE. Limitations of current risk-adjustment models in the era of coronary stenting. *Am Heart J* 2003; 145: 683-92
3. Nashef SAM, Roques F, Michel P, Gauducheau E, Lemeshow S, Salomon R, the EuroSCORE studt group. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardio-thorac Surg* 1999;16:9-13
4. Roques F, Nashef SAM, Pinna Pintor P, David M, Baudet E. Does EuroSCORE work in individual European countries? *Eur J Cardio-thorac Surg* 2000;18:27-30
5. Pons JMV, Granados A, Espinas JA, Borrás JM, Martín I, Moreno V. Assessing open heart surgery mortality in Catalonia (Spain) through a predictive risk model. *Eur J Cardio-thorac Surg* 1997; 11: 415-423
6. Roques F, Michel P, Goldstone AR, Nashef SAM. The logistic EuroSCORE. *Eur Heart J* 2003;24:1-2
7. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing areas under two or more correlated receiver operating curves: a nonparametric approach. *Biometrics* 1988;44:837-845
8. DeLong ER, Peterson ED, DeLong DM, Muhlbaier LH, Hackett S, Mark DB. Comparing risk-adjustment methods for provider profiling. *Stat Med* 1997;16:2645-2664
9. Steyerberg EW, Eijkemans MJC, Boersma E, Habbema JDF. Equally valid models gave divergent predictions for mortality in acute myocardial infarction patients in a comparison of logical regression models. *J Clin Epidemiol* 2005; 58: 383-390
10. Hamley J. In: McNeil BJ, editor. The meaning and use of the area under receiver operating characteristic (ROC) curve. *Radiology* 1992;143:29 – 36
11. Ivanov J, Tu JV, Naylor CD. Ready-made, recalibrated, or remodeled? Issues in the use of risk indexes for assessing mortality after coronary artery bypass graft surgery. *Circulation* 1999;99:2098-2104
12. Peterson ED, DeLong ER, Muhlbaier LH, Rosen AB, Buell HE, Kiefe CI, Kresovik TF. Challenges in comparing risk-adjusted bypass surgery mortality results. *JACC* 2000;36:2174-84
13. Hannan EL, Wu C. Assessing quality and outcomes for percutaneous coronary intervention: choosing statistical models, outcomes, time periods, and patient populations. *Am Heart J* 2003;145:571-574
14. García-Fuster R, Montero JA, Gil O, Hornero F, Cánovas S, Bueno M, Buendía J, Rodríguez I. Trends in coronary artery bypass surgery: Changing type of surgical patient. *Rev Esp Cardiol* 2005;58:512-522
15. Pons JMV, Moreno V, Borrás J, Espinas J, Almazan C, Granados A. Open heart surgery in public and private practice. *J Health Serv Res Policy* 1999;4:73-78

16. Permanyer-Miralda G, Brotons C, Ribera A, Alonso J, Cascant P, Moral I and the CIRCORCA study group. Outcomes of coronary artery surgery: Determinants of quality of live related to postoperative health. *Rev Esp Cardiol* 2001; 54: 607-616
17. Pons JMV, Borrás JM, Espinas JA, Moreno V, Cardona M, Granados A. subjective versus statistical assessment of mortality risk in open heart surgical procedures. *Ann thorac Surg* 1999;67:635-40
18. Grover FL, Shroyer AL, Hammermeister K, Edwards FH, Ferguson TB, Dziuban SW, Cleveland JC, Clark RE, McDonald G. A decade's experience with quality improvement in cardiac surgery using the Veterans Affairs and Society of Thoracic Surgeons national databases. *Ann Surg* 2001;234:464-72
19. Shahian DM, Blackstone EH, Edwards FH, Grover FL, Grunkemeier GL, Naftel DC, Nashef SA, Nugent WC, Peterson ED; STS workforce on evidence-based surgery. Cardiac surgery risk models: a position article. *Ann Thorac Surg* 2004;78:1868-77

Figure legends

Figure 1. Predictive ability for inhospital mortality of the CAHTA risk score and the EuroSCORE.

Figure 2. Top, plot of the expected number of deaths vs observed number of deaths in each decile of the estimated probability using the CAHTA risk score. Bottom, plot of the expected number of deaths vs the observed number of deaths in each decile of the estimated probability using the EuroSCORE. Dashed lines represent a perfect fit between observed and expected values (intercept=0, slope=1).

Figure 3. Adjusted mortality (observed / expected) of the participating hospitals using the CAHTA model and the EuroSCORE, before and after recalibration.

Figures

Figure 1

	CAHTA model	EuroSCORE
Area under the ROC curve	0.74 (0.67 – 0.81)	0.76 (0.70 – 0.82)

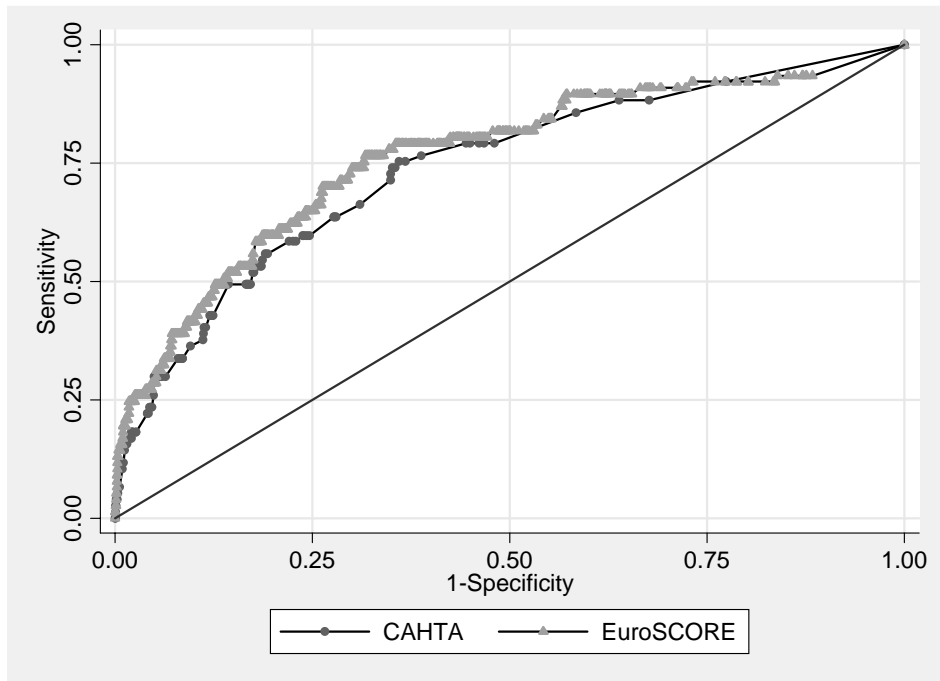


Figure 2.

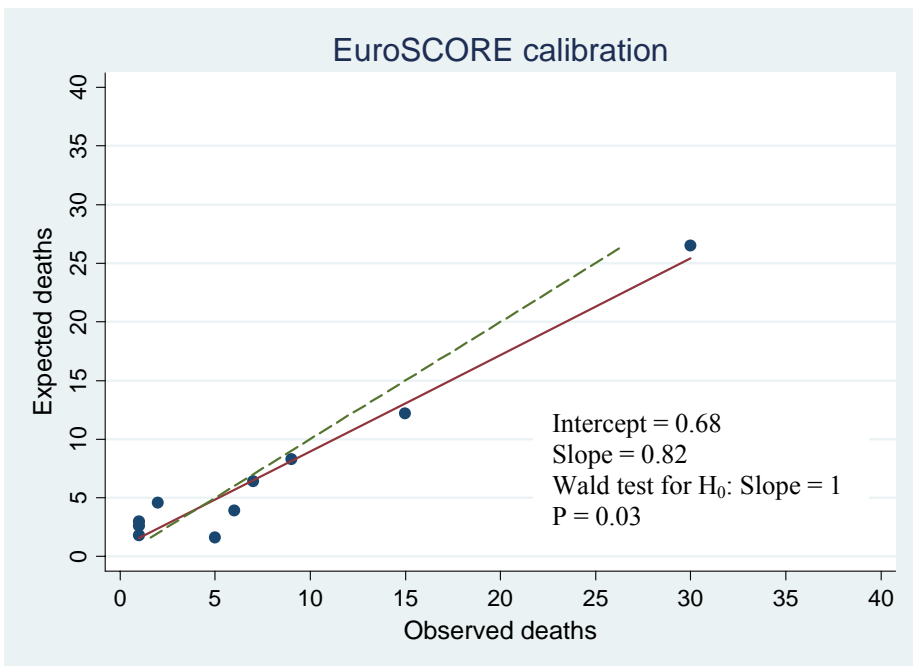
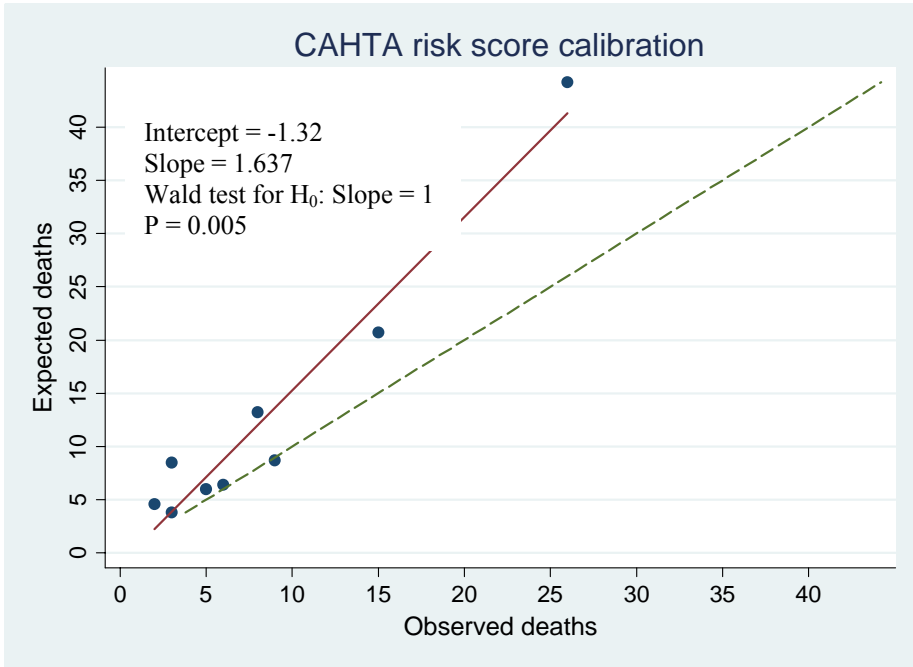
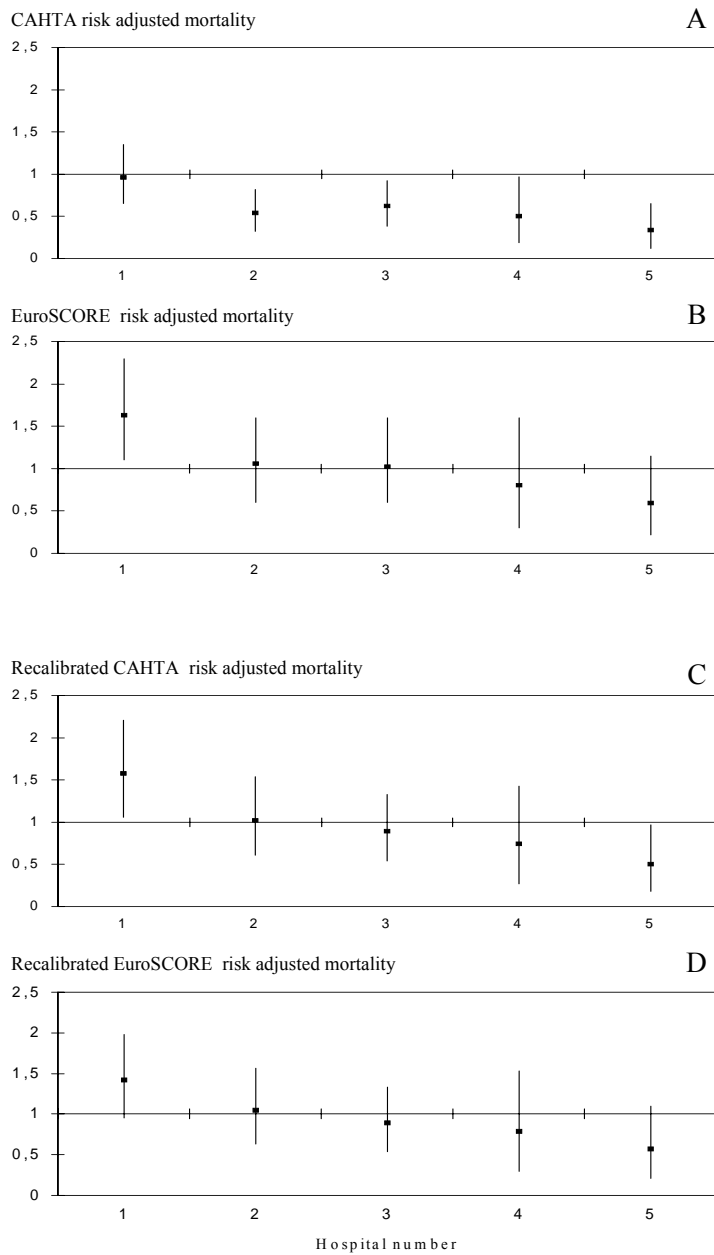


Figure 3.



Tables

Table I. OR for inhospital mortality of the CAHTA risk factors in the development sample and in the revalidation sample.

CAHTA risk factors	Score weight	Prevalence in the development sample (n=583) *	OR in the development sample	Prevalence in the revalidation sample (n=1,605)	OR (95% CI) in the revalidation sample
Age					
70 – 79	7	17.3 %	2.1	36.2 %	2.3 (1.4 – 3.8)
≥ 80	17	1 %	5.2	2.1 %	2.4 (0.7 – 8.7)
Recent myocardial infarction (<29 days)	10	14.8 %	2.6	19.5 %	1.6 (0.9 – 2.8)
Functional class					
III	4	46.3 %	1.5	23.8 %	1.4 (0.8 – 2.6)
IV	10	26.1 %	2.6	13.3 %	2.6 (1.4 – 5)
Liver disease	8	1.9 %	2.3	3.1 %	0.4 (0.1 – 2.9)
Left ventricular aneurysm	11	4.6 %	3	3.4 %	2.1 (0.7 – 5.8)
Creatinine ≥ 1.5 mg/dl	8	10.3 %	2.3	10.1 %	3.5 (2 – 6)
Cardiogenic shock	13	0.9 %	3.7	1 %	3.1 (0.5 – 18.4)
Presurgical mechanical ventilation	7	0.7 %	2	0.5 %	2.4 (0.3 – 20.1)
Priority (emergent/urgent)	4	49.1 %	1.4	53 %	1.3 (0.7 – 2.4)
Reoperation					
First	9	-	2.4	-	-
Second	15	-	4.3	-	-
Mitral valve surgery	6	-	1.8	-	-
Tricuspid valve surgery	10	-	2.7	-	-
Thoracic aortic surgery	12	-	3.5	-	-
Combined valve and coronary surgery	7	-	2	-	-
Mean CAHTA risk score (95% CI)		9.45 (8.7 – 10.2)		9.47 (9 – 9.9)	
Inhospital mortality (95% CI)		7.9 (5.8 – 10.4)		4.8 (3.8 – 6)	

*Excluding reoperations and valve or combined valve and coronary procedures

Table 2. OR for in-hospital mortality of the EuroSCORE risk factors in the development sample and in the revalidation sample.

EuroSCORE risk factors	Score weight	OR in development sample	Prevalence in the study population	OR in revalidation sample
Age (mean, SD)	1	1.1	64.8 (9.9)	1.05 (1 – 1.1)
Female	1	1.4	20 %	1 (0.5 – 1.9)
Chronic pulmonary disease	1	1.6	16 %	2.5 (1.5 – 4.4)
Extracardiac arteriopathy	2	1.9	25 %	1.7 (1.04 – 2.9)
Neurological dysfunction	2	2.3	2.4 %	0.6 (0.1 – 4.5)
Previous cardiac surgery	3	2.6	-	-
Serum creatinine > 200 µmol/l	2	1.9	10 %	2.7 (1.6 – 4.8)
Active endocarditis	3	2.5	-	-
Critical preoperative state	3	2.2	3 %	2.3 (0.9 – 5.9)
Unstable angina (need for intravenous nitrates)	2	1.5	12 %	2.7 (1.5 – 4.9)
LV disfunction				
Moderate	1	1.5	22 %	1.01 (0.6 – 1.8)
Severe	3	2.5	4 %	3.1 (1.4 – 6.9)
Recent AMI (<90 days)	2	1.6	26 %	1.4 (0.8 – 2.4)
Pulmonary hypertension	2	2	1 %	4.9 (0.9 – 26.3)
Emergency	2	2.8	1.2 %	2.1 (0.5 – 8.3)
Other than isolated CABG	2	1.6	-	-
Surgery on thoracic aorta	3	3.2	-	-
Postinfarct septal rupture	4	3.8	0.1 %	-
Mean EuroSCORE			4.03	

ARTICLE 3

Ferreira-González I, Ribera A, Cascant P, Permanyer-Miralda G for the ARCA study group. Outcomes in off-pump versus on-pump coronary artery bypass grafting stratified by preoperative risk profile: an assessment using propensity score. *Eur Heart J* 2006;27:2473-80

Ferreira-Gonzalez I, Ribera A, Cascant P, Permanyer-Miralda G.
(grupo de estudio ARCA)

*Outcomes in off-pump versus on-pump coronary artery bypass
grafting stratified by operative risk profile: an assessment using
propensity score*

European Heart Journal 2006 Oct;27(20):2473-80. Epub 2006
Sep 15.

DISCUSSIÓ

L'objectiu principal d'aquest estudi era comprovar si el tipus de gestió hospitalària (pública o privada) s'associava a la sanitat pública catalana amb la taxa de mortalitat hospitalària després de la cirurgia de derivació aortocoronària. Aquesta hipòtesi es va originar fa més de deu anys, quan els resultats de la cirurgia coronària en el nostre àmbit eren clarament inferiors als desitjables. La sospita de que aquest mal resultat estava associat al procés assistencial va motivar aleshores un interès particular en avaluar els resultats de la cirurgia cardíaca en general i particularment de la cirurgia de derivació aortocoronària (Pons et al, 1999 i Permanyer et al, 1998).

Alhora que aquest estudi ens donava l'oportunitat d'avaluar els resultats quirúrgics locals, ens ha permès: Primer, contribuir al coneixement de la metodologia de l'avaluació del risc, cada cop més complexa i necessària en la medicina contemporània; i segon, analitzar l'efectivitat en el nostre medi d'una tècnica quirúrgica de recent introducció amb una estratificació adequada de la població segons el risc preoperatori.

El primer estudi i el segon son complementaris i val la pena discutir-los de manera conjunta. En el primer article es presenten les troballes més importants pel que fa a l'objectiu principal de l'estudi, és a dir, comprovar si el tipus de gestió hospitalària (pública o privada) s'associa a la sanitat pública catalana amb una taxa de mortalitat hospitalària diferent. En el segon es fa un anàlisi minuciós dels dos instruments d'avaluació del risc utilitzats. Més enllà de que els resultats d'aquesta segona anàlisi puguin ser, en part, generalitzables a altres àmbits i a altres models de risc, l'estudi crític i minuciós dels avantatges i limitacions de les escales de risc és necessari per a interpretar correctament les troballes del primer. De manera que per poder interpretar el resultat de l'avaluació correctament es necessari discutir abans dos aspectes importants de l'anàlisi: 1) Fins a quin punt és vàlid l'ajust del risc realitzat?; i 2) Podem atribuir la diferència que, encara que petita, hem observat en els resultats ajustats entre centres públics i centres privats, a un diferent procés assistencial segons el tipus de gestió?

El tercer estudi, que s'ha dut a terme en la mateixa cohort de pacients que els dos precedents, representa un exemple d'una de les aplicacions que poden tenir les escales de risc en els estudis d'efectivitat en la pràctica clínica real en un context determinat.

Estructurarem la discussió de la següent manera: discutirem, en primer lloc, quines son les limitacions de les escales de risc en general i particularment quina és la validesa dels dos instruments utilitzats a l'hora d'avaluar els resultats de la cirurgia coronària en el nostre medi; en segon lloc, quines consideracions cal fer abans d'atribuir les diferències observades al procés assistencial; i en tercer lloc, comentarem les troballes de l'estudi de

l'efectivitat de la cirurgia sense CEC. Finalment, comentarem quines son les limitacions, les implicacions de tot plegat i les necessitats de recerca futura.

Les escales de risc

Son ben conegudes les limitacions que poden tenir els sistemes d'ajust del risc (Cortina-Romero JM, 2005 i Weintraub WS, 2006) però val la pena explicar-les aquí amb prou detall.

El problema dels factors de risc rellevants

Les escales de risc no poden incloure tots els factors de risc potencialment importants. Algunes característiques que els professionals, clínics i cirurgians, considerarien importants en funció de la seva pròpia experiència per valorar el risc del pacient individual no solen tenir-se en compte en les models de risc. Això s'il·lustra de manera molt clara en un treball de la sèrie de l'estudi de l'AATRM (Pons JMV et al, 1999), en el que es comparava l'escala de risc desenvolupada a partir del model estadístic amb la predicció del risc segons la percepció subjectiva del cirurgià. Aquesta comparació va ser útil en aquell moment per validar l'escala que havien desenvolupat perquè globalment el model estadístic presentava millor precisió en la predicció del risc, sobretot en els estrats de risc intermedi on, d'altra banda, s'acumulen la gran majoria dels casos. Però es feia una observació interessant, i és que la predicció subjectiva per part dels cirurgians era més precisa en el dos extrems de risc. Aquest resultat s'interpretà pel fet que l'avaluació subjectiva inclou aquests elements infreqüents o difícils de quantificar de manera objectiva.

Algunes situacions com, per exemple, si el pacient està en tractament amb agents immunosupressors, jugarien un paper clau en la decisió mèdica però no apareixen en les escales de risc senzillament perquè son poc habituals. L'omissió de característiques infreqüents també té relació amb la grandària de la mostra que ha servit per a la derivació del model de risc. En una població gran en la qual hi hagi prou casuística fins i tot de pacients amb condicions poc habituals, aquestes característiques podrien aparèixer amb una associació significativa amb el resultat. Per exemple en els models que es generen a partir de les bases de dades americanes com les de l'estat de Nova York o de la STS (Hannan EL, 2006 i Shroyer ALW, 2003) s'inclouen variables que difícilment trobem en les escales desenvolupades en zones geogràfiques d'extensió més limitada. En el nostre cas, quan comparem els factors de risc que apareixen a l'escala de l'AATRM i els que apareixen a l'EuroSCORE (Taules 1 i 2 de l'article 2) hi observem alguns factors com la disfunció neurològica, l'estat crític preoperatori i la hipertensió pulmonar que, encara que s'haguessin registrat en la mostra de poc menys d'un miler d'intervencions utilitzades per derivar l'escala de l'AATRM, és poc probable

que hi fossin en una prevalença suficient com per aparèixer en l'equació resultant.

D'altra banda, utilitzar escales externes pot suposar també la omissió d'informació que pot ser important en el context local. Per exemple, la presència de malaltia pulmonar obstructiva crònica, molt prevalent en la nostra població i molt important a l'hora de determinar el risc quirúrgic, no apareix en la predicció del risc en altres poblacions del nord d'Amèrica com Ontario (Tu JV et al, 1995), però sí que apareix en la recent escala desenvolupada a l'estat de Nova York (Hannan et al, 2006). En canvi, en aquesta última no s'hi inclou la indicació urgent o algun indicador d'instabilitat clínica, que és un fort predictor de mortalitat en el nostre medi. Potser aquesta omissió es deu a una voluntat explícita de no incloure en el model variables relacionades amb el procés assistencial per tal de que aquest sigui més fàcilment exportable però, sens dubte, suposaria una important infraestimació del risc si s'apliqués l'escala en el nostre context local. En aquest sentit, és remarcable la diferència que observem en al OR de mortalitat hospitalària per l'angina inestable entre la mostra que es va utilitzar per derivar l'EuroSCORE i la població de l'estudi.

També és habitual que els professionals que han de prendre decisions sobre el malalt individual reivindiquin que els models estadístics no tinguin en compte situacions relacionades amb l'estat general del pacient o el que s'anomena la *fragilitat*, poc susceptible, encara, de mesura objectiva i, en canvi, ben percebuda de forma intuïtiva per clínics experts. Aquí entrariem de ple en la discussió sobre la necessitat d'objectivitat a l'hora de definir les variables que intervenen en un model. És necessari que la definició d'un model de risc sigui el màxim d'objectiva possible, reproducible i exportable, és a dir, que a l'hora d'aplicar el model en una població diferent de la que es va derivar sigui possible obtenir la informació de manera objectiva i amb idèntica significació que en la població original.

És important destacar que, tant en l'escala de l'AATRM com a l'EuroSCORE, trobem factors de risc que poden suposar-se o bé de definició massa subjectiva o bé poc exportables: la classificació del grau funcional i de la indicació electiva o urgent en la primera i la necessitat de nitrats intravenosos en la segona; tots ells indicadors d'instabilitat de la malaltia. La classificació del grau funcional, tot i que té una definició clara i precisa, està sotmesa a la subjectivitat de l'investigador i al fet que no queda mai clar a quin moment es refereix (al grau màxim assolit, en el moment de l'ingrés o en el moment de la indicació quirúrgica, etc.). D'altra banda, el criteri d'urgència o de necessitat de nitrats intravenosos pot dependre del procés assistencial i dels diferents hàbits terapèutics de cada centre particular. Nosaltres vam incloure en el protocol d'estudi una altra definició de situació inestable (haver tingut una crisi d'angina de repòs en les 72 hores prèvies a la

intervenció), precisament per validar l'objectivitat d'aquests indicadors d'instabilitat presents a les escales de risc.

A la vista dels resultats (major nombre de malalts amb classe funcional IV, amb indicació urgent o emergent de la cirurgia i amb necessitat de nitrats intravenosos en els centres de gestió privada), cal suposar que en els hospitals privats analitzats, la política assistencial permet o condueix a una operació més precoç dels pacients inestables. S'accepta, doncs la majoria d'escales de risc ho contemplen, que una operació més propera a la fase d'instabilitat (angina inestable) comporta un major risc quirúrgic. Si bé pot argumentar-se que la diferència en l'ús de nitrats intravenosos pot reflectir diferents hàbits terapèutics més que a una autèntica diferència en el risc basal, aquest risc més elevat queda il·lustrat pel fet que, independentment del tipus de gestió del centre, la operació de pacients amb necessitat de nitrats intravenosos era també predictor de major mortalitat (OR de 2,4). El mateix succeïa si en comptes de la necessitat de nitrats intravenosos incloïem el fet d'haver tingut una crisi d'angina de repòs en les 72h prèvies a la intervenció. Per tant, a part de la influència d'hàbits assistencials, podem suposar sòlidament que els malalts operats en centres privats tenien un risc quirúrgic més elevat.

De condicions potencialment relacionades amb el risc quirúrgic n'hi ha, doncs, moltes, i ens podem qüestionar si les escales, degut als requeriments de parsimònia que tenen els models estadístics, no es queden curtes a l'hora de tenir-los en compte tots i cada un. Sobre això cal afegir que la parsimònia no és només un requeriment estadístic, també és un requeriment pràctic a l'hora de fer un càlcul senzill de la probabilitat d'un mal resultat a partir d'una determinada combinació de factors de risc. Però el principi de parsimònia no significa tant sols economitzar l'esforç en la recollida i codificació de dades sinó que implica trobar la combinació òptima entre l'esforç mínim i el màxim d'informació continguda en el conjunt de dades.

Està ben establert que amb la informació continguda en cinc o sis factors s'obté prou precisió en la predicció del risc i que aquesta capacitat de predicció no augmenta significativament a partir d'aquest nombre de variables (Tu JV et al, 1997). En la nostra experiència hem pogut comprovar com utilitzant tres sistemes diferents per ajustar pel risc (dues escales de risc i un model de regressió logística elaborat a partir de les pròpies dades), tots tres amb una combinació de factors diferent, obtenim un resultat pràcticament equivalent en l'avaluació. D'altra banda, observem també uns trets comuns (l'edat, la cirurgia prèvia, la severitat del símptomes en el preoperatori i la presència de certes malalties cròniques) que es repeteixen també en la resta de models de predicció de la mortalitat quirúrgica que apareixen a la literatura.

L'equivalència en els resultats utilitzant diferents estratègies malgrat les diferències i la presència gairebé constant de certes característiques comunes en totes les escales de predicció del risc indiquen que la sola inclusió d'aquestes característiques fonamentals seria vàlida per a un correcte ajust del risc en la majoria de casos.

El problema de la població de referència i la seva influència en la validesa dels models: EuroSCORE vs escala local

Un altre tema important, sovint debatut però escassament analitzat (Peterson ED et al, 2000), és el de la població de referència. La validesa de les dues poblacions que van servir per a la construcció dels dos models, com a referència per a la nostra població d'interès, és un problema fonamental.

Els dos models de risc utilitzats en aquesta avaluació necessitaven revalidar-se per raons diferents. L'EuroSCORE perquè l'aplicàvem en un context inclòs en el context més gran per al qual va ser derivat però amb unes característiques pròpies que poden ser diferents de les característiques globals de la població europea. I l'escala de l'AATRM perquè tot i que es va desenvolupar específicament per al context en el qual l'apliquem, han passat deu anys des de la seva derivació. Per això ens vam proposar estudiar-ne la validesa i al mateix temps comparar-ne les propietats amb la idea d'estudiar els avantatges i limitacions d'una aproximació sobre l'altra.

Per determinar la validesa de les dues escales vam estudiar-les des de dos punts de vista: primer, des del punt de vista de les seves propietats estadístiques de discriminació i calibració; i segon des del punt de vista del seu funcionament en l'estudi d'avaluació, és a dir, estudiant les diferències entre l'una i l'altra a l'hora d'estimar la mortalitat ajustada.

La discriminació indica la capacitat del model per diferenciar entre els pacients que moren i els que no. La calibració es refereix al grau de concordança entre la predicció de risc del model i la incidència real de mortalitat en tots els estrats de risc.

Ambdues escales tenien en l'estudi de la discriminació una àrea sota la corba ROC d'aproximadament 0,75. Aquesta no és pas una capacitat de predicció òptima però, no obstant similar a la que s'obtenia en els estudis originals de validació. La calibració en canvi era subòptima i, sobretot l'escala de l'AATRM, tenia una important tendència a sobreestimar el risc. En fer l'exercici de recalibració interna es va obtenir un ajust perfecte entre els valors observats i els valors predits per la nova escala recalibrada.

Era fàcil anticipar que el model de l'AATRM, construït 10 anys enrere en una època en la qual hi havia raons per suposar que els resultats de la cirurgia

eren molt pitjors als del moment en que es recollien les dades per al present estudi, tindria una clara tendència a sobreestimar el risc.

Pel que fa a l'EuroSCORE, ens podíem preguntar fins a quin punt un instrument derivat d'un conjunt de poblacions europees probablement amb característiques divergents entre elles, pot representar una referència per als nostres resultats locals. Ja s'ha remarcat la diferència que en aquell moment hi havia entre els resultats a Espanya i els resultats en altres països europeus (cita EuroSCORE països), essent la mortalitat observada en els hospitals espanyols que participaven del 6,8%, molt diferent a l'1,5% de Finlàndia o al 3,7% del Regne Unit, i també lluny de la d'altres països mediterranis com Itàlia (2,4%) o França (3,2%). Aquesta diferència assenyalava llavors la mortalitat quirúrgica a Espanya com a significativament superior a l'esperada, fet probablement explicat per una qualitat assistencial realment pitjor que en altres països europeus més que per un mal funcionament de l'escala. Així doncs, suposant una millora en els resultats locals en els últims anys fins a assolir els nivells que s'observaven aleshores a Europa no era estrany esperar un millor ajust de l'EuroSCORE a les dades actuals.

La capacitat avaluativa (o la influència del model aplicat en l'estimació del risc ajustat dels centres participants) es va determinar comparant qualitativament els resultats obtinguts amb les dues escales: Com es distribuïen els centres en funció de la seva mortalitat ajustada i quins havien de ser considerats *outliers* en cada cas. Quan no existeix un criteri de referència ideal les comparacions son sempre relatives i es tendeix a considerar el terme mitjà de la realitat observada com a referència. És a dir, considerem aquí com a més vàlid aquell model que més s'aproxima al resultat observat perquè no tenim manera de determinar quin és el resultat ideal assolible.

Vam observar que, malgrat que l'escala de l'AATRM sobreestimava el risc, l'ordenació dels centres segons la mortalitat ajustada era idèntica utilitzant els dos models. La detecció d'*outliers* en canvi estava molt influïda per la capacitat de calibració: abans de recalibrar-la, l'escala de l'AATRM identificava quatre dels cinc centres estudiats com a *outliers* en el sentit de millor resultat de l'esperat. Aquesta troballa reforça la idea, ja observada per altres (Hannan EL i Wu C, 2003), de que els models antics o desenvolupats en un context diferent del que s'apliquen, poden funcionar raonablement bé sempre que no s'utilitzin per identificar *outliers*. Quan el que es pretén és detectar quines institucions o equips quirúrgics tenen un resultat per sobre o per sota del que seria esperable, cal disposar d'escales ben calibrades o bé aplicar una referència interna.

Aquests resultats son consistents amb els d'altres estudis (Peterson ED et al, 2000), en els que s'observa que: malgrat les diferències importants en la capacitat de calibració i la tendència a sobreestimar el risc (sobretot en els estrats de risc alt i en escales més antigues), la comparació interna entre

centres depèn poc de l'escala utilitzada, mentre que la identificació d'*outliers* varia de manera important en funció de la referència utilitzada.

La possibilitat de recalibració

En aquest estudi es va aplicar una estratègia de recalibració (recalibració interna o, com l'anomenen alguns autors, "recalibració a l'engròs") (Peterson ED et al, 2000 i Steyerberg et al, 2005) que consisteix en ajustar un nou model de regressió logística sobre la puntuació de l'escala (no sobre els factors de risc individuals) i estimar així uns nous termes α i β de manera que la nova probabilitat (mortalitat esperada) pot calcular-se de la següent manera:

$$P = \exp(\alpha + \beta R) / [1 + \exp(\alpha + \beta R)]; \quad R: \text{Puntuació de Risc}$$

Hi ha altres estratègies de recalibració possibles. Una possibilitat és recalcular de nou els coeficients de tots els factors de risc (Ivanov J et al, 1999) però això requereix una mostra prou gran per poder dividir-la en una cohort de derivació i una de validació i obtenir estimacions prou precises dels coeficients. Una altra possibilitat més senzilla i que no necessita cap tècnica d'anàlisi complexa seria corregir la mortalitat estimada segons el diferencial de mortalitat observada entre la cohort de derivació i la cohort d'estudi de la següent manera (Hannan EL et al, 2006):

$$(MA)_e = (MA)_d \times (MO)_e / (ME)_e$$

$(MA)_e$ = Mortalitat ajustada segons el risc en la població d'estudi

$(MA)_d$ = Mortalitat ajustada segons el risc en la cohort de derivació

$(MO)_e$ = Mortalitat observada en la població d'estudi

$(ME)_e$ = Mortalitat esperada en la població d'estudi basada en el model

En tots els casos la recalibració significa que el criteri de referència passa a ser intern i els resultats entre proveïdors es comparen amb un valor mitjà entre tots ells. Això té també les seves limitacions, doncs en el cas que els resultats en el context estudiat fossin globalment molt allunyats del que seria esperable no ho detectaríem un cop recalibrada l'escala. En canvi, funciona a l'hora de detectar vertaders *outliers* i cridar l'atenció sobre possibles problemes de qualitat assistencial.

Un avantatge que té el model de L'AATRM sobre l'EuroSCORE és que es basa en una mostra exhaustiva de la població a la qual es pretén utilitzar i, per tant, és representatiu de la situació en un context determinat i en un moment en el temps. Això el fa directament aplicable en aquest context, obviant el problema de si la població de referència és adequada. Aquest fet i no cap altre ens permet, en el nostre cas, interpretar de manera inequívoca

la pèrdua de calibració com a un canvi real en el resultat en els deu anys que han passat des de la seva derivació fins a la seva aplicació actual. Així doncs, en determinades condicions (quan el model era correcte en origen i la població és la mateixa), la falta de calibració és el reflex d'un canvi temporal en els resultats més que no pas un defecte del model (Krumholz HM et al, 1999).

És per tant discutible fins a quin punt és necessària la recalibració i en quines situacions. Els models basats en dades antigues d'un determinat sistema sanitari tindran tendència amb el temps a sobreestimar el risc de manera global, però això no afecta necessàriament la seva utilitat com a eina per a comparar el resultat de manera relativa entre centres i per a estimar, a més, la magnitud dels canvis que s'han produït en el temps, sempre i quant la falta de calibració es produeixi en tots els estats de risc de manera homogènia. En la nostra experiència, la tendència a la sobreestimació es produeix sobretot en els estrats de risc més alt, per tant afavorint (en el sentit d'estimar mortalitats esperades més altes respecte les observades) els grups amb puntuacions extremes de l'escala de l'AATRM.

Pel que fa a l'EuroSCORE, hem comprovat que, segons els criteris habituals de discriminació i calibració, la seva validesa per a la població estudiada és semblant a la que tenia en el moment del seu desenvolupament en aplicar-lo a la població espanyola.

Seria per tant exagerat jutjar la validesa d'un model només per les seves propietats estadístiques de discriminació i calibració, doncs, malgrat que un bon model ha de tenir bones propietats en el moment del seu desenvolupament, la falta de calibració en una aplicació local pot indicar o bé un canvi temporal en els resultats quan utilitzem un model local o bé una diferència en els resultats respecte un context global quan utilitzem un model d'àmbit general. En el cas que ens interressi tenir un criteri de referència intern de comparació (per exemple si volem detectar possibles *outliers* en un determinat context) n'hi hauria prou amb aplicar un mètode de recalibració senzill.

Influència del tipus de gestió

Hem estudiat la validesa de les escales utilitzades i sabem que: ambdues tenen una capacitat de discriminació acceptable, és a dir, prediuen correctament en un elevat percentatge de casos la mortalitat d'un pacient a partir de la seva particular combinació de factors de risc; i que el fet que l'escala de l'AATRM sobreestimi el risc només té implicacions a l'hora de detectar *outliers* (problema que es resol fàcilment mitjançant la recalibració interna). Hem vist també com l'aplicació dels dos models diferents d'ajust del risc (un local i un europeu) mostra resultats molt semblants tant si fem una

anàlisi estratificada: la OR per a mort hospitalària dels centres privats respecte als públics era 0,56 (IC 95%: 0,29 – 1,06) utilitzant l'EuroSCORE i 0,56 (IC 95%: 0,29 – 1,07) utilitzant l'escala de l'AATRM; com si comparem les mortalitats ajustades entre els centres agrupats segons el tipus de gestió. Si fem un anàlisi alternatiu ajustant el resultat per les característiques basals dels pacients utilitzant un model de regressió logística obtenim també un resultat equivalent: OR de 0,43 (IC 95%: 0,21 – 0,89). Tot això ens permet assegurar que l'ajust realitzat és vàlid i que, per tant, hi ha una diferència de significació estadística marginal però de magnitud no menyspreable, en la mortalitat hospitalària entre centres de gestió pública i centres de gestió privada.

Ara bé, si aquesta diferència és atribuïble al tipus de gestió, és a dir, si les diferències en el procés assistencial entre els dos tipus de gestió s'associen a un resultat diferent, s'ha d'interpretar a la llum de certes consideracions. És important assenyalar que la naturalesa d'aquesta associació hauria de ser objecte d'estudis més específics, segurament amb l'ús de metodologies que permetessin l'estudi de processos i mecanismes complexos.

La primera consideració que cal fer és si el resultat pot associar-se a l'efecte individual dels centres més que al fet d'agrupar-los segons el seu tipus de gestió, ja que els cinc centres participants tenen característiques pròpies més enllà del fet d'estar gestionats per un organisme públic o privat. Els dos centres privats són els únics a Catalunya en els quals s'operen malalts procedents de la sanitat pública. El Centre 5 és un centre especialitzat en malaltia cardiovascular i rep malalts de la sanitat pública per operar-se quan el processos habituals en els centres públics de referència se saturen. El Centre 4 en canvi és un hospital general que a través de concert amb l'Institut Català de la Salut constitueix centre de referència per una àrea geogràfica determinada. Els tres centres públics són tots grans hospitals docents de referència amb característiques d'estructura similars entre ells.

Certament, les característiques d'estructura i del procés assistencial de cada centre podrien explicar en part els resultats observats (González-Santos JM, 2006). González Santos descriu en l'editorial (vegeu-ne una còpia a l'apartat Annex 4) sobre aquest treball algunes de les diferències entre els centres de gestió pública i privada que bé poden aplicar-se als centres del sistema públic català. Aquestes diferències, que també es descriuen breument en la discussió de l'article, es refereixen a característiques de l'estructura de les institucions proveïdores de l'atenció sanitària de les quals se n'ha estudiat la influència sobre la qualitat assistencial.

De manera remarcable, quan introduïem en el model logístic una variable indicadora de l'activitat del cirurgià (nombre d'intervencions realitzades per cada cirurgia en la població de l'estudi), aquesta apareixia també associada a una major supervivència però no modificava significativament l'estimació dels

paràmetres de les altres variables incloses en el model. L'associació entre el tipus de gestió i la mortalitat hospitalària disminueix lleugerament passant la OR per al caràcter de gestió privada de 0,43 a 0,49 (IC 95%: 0,24 – 1), indicant que l'activitat del cirurgià pot ser, entre molts d'altres, un dels factors que expliquin la diferència entre centres públics i privats. És ben plausible que els diferents models de gestió determinin uns patrons d'activitat dels cirurgians diferents, com suggereix el conjunt de la nostra experiència (CIRCORCA, ARCA), en la que trobem una taxa d'intervencions per cirurgia més favorable en els centres de gestió privada.

Els treballs en els quals s'estudia l'efecte del volum d'intervencions per cirurgia i centre disposen de bases de dades enormes. És el cas del registre clínic de l'estat de Nova York (Hannan EL et al, 2003), amb dotzenes de milers d'intervencions, y la iniciativa de la *Society of Thoracic Surgeons* (Peterson ED et al, 2004) que incorpora anualment la informació sobre el risc i el resultat de més de 160.000 intervencions d'empelt coronari aïllat. En ambdós casos el gran nombre de centres i cirurgians participants permet realitzar una anàlisi precisa de la qualitat assistencial incorporant característiques dels proveïdors a més del perfil dels pacients.

Malauradament és difícil realitzar estudis com aquests a Catalunya. Primer, perquè es necessiten una gran quantitat de recursos i d'esforç per a realitzar registres de magnitud i qualitat suficient per a aquest propòsit. I segon, perquè en l'avaluació limitada a una sola comunitat autònoma, el nombre de centres és petit i relativament homogeni quant al volum assistencial i a altres característiques (com el caràcter docent i el nivell d'especialització), d'altra banda, difícilment dissociables del model de gestió. De manera similar, el caràcter docent dels centres públics pot fer-los més proclius a desigualtats en la qualitat assistencial (khuri SF et al, 2001).

Les característiques d'estructura descrites juntament amb altres peculiaritats de cada centre, com per exemple les diferents pràctiques en les cures postoperatòries, podrien explicar les troballes d'aquest estudi, però és important reconèixer que aquestes no son necessàriament dissociables del tipus de gestió. Igual que passa amb el caràcter docent o el nombre d'intervencions per cirurgia (més eficient en els centres de gestió privada) es pot intuir que el caràcter públic o privat engloba diverses característiques diferencials, de vegades difícils d'observar, que poden associar-se al resultat del procediment. Cal tenir en compte a més que no es van trobar diferències significatives entre centres, ni en el resultat, ni en els factors de risc més importants i tampoc es va poder demostrar una influència del centre en un model multinivell. En canvi, sí que observem una tendència a que tant les característiques de risc com el resultat ajustat siguin homogènies entre els centres dins de cada grup definit segons el tipus de gestió. És a dir, en aquelles característiques que els diferencien (malaltia de tres vasos, afectació de la DA proximal, cirurgia urgent o emergent, angina inestable, creatinina

elevada, MPOC, estat preoperatori crític, disfunció ventricular esquerra), s'observa una tendència a que els tres centres públics es situïn en un extrem i els dos privats a l'altre, igual que quan ordenem els centres segons la mortalitat ajustada. Per tant, aquest estudi suggereix clarament que el tipus de gestió, ni que sigui indirectament a través d'altres variables, pot exercir una influència sobre la qualitat dels resultats. No obstant, per avaluar exactament les diferències en la qualitat assistencial o els problemes concrets que expliquen aquest fet caldria fer estudis més específics o auditories individuals.

Millora dels resultats associada als canvis en el procés assistencial

Una troballa important d'aquest estudi és el fet d'haver constatat una millora important dels resultats de la cirurgia coronària a la sanitat pública de Catalunya en els últims anys. La constatació més clara és la que s'evidencia quan es comparen alguns dels indicadors de risc més importants i la mortalitat observada entre el CIRCOCA (1997) i l'ARCA (2001-2002), perquè ambdós treballs es van realitzar en els mateixos centres, aplicant els mateixos criteris d'inclusió i idèntica definició de les variables. La mortalitat es redueix a la meitat (del 10% al 4,9%) amb una prevalença dels indicadors de risc més importants equivalent, si no superior, en la cohort analitzada a l'estudi ARCA.

De la comparació que es fa en el segon article entre el perfil de risc i la mortalitat hospitalària entre la cohort de derivació (un cop excloses les reintervencions i les intervencions valvulars o combinades) i la població d'estudi, es pot extraure la mateixa conclusió. La mortalitat es redueix del 7,9% al 4,8% mentre que el risc, calculat segons la mateixa equació, és equivalent (la puntuació de l'AAATRM passa de 9,45 a 9,47).

Les dues comparacions apunten a una reducció de la mortalitat, a igualtat de risc, entre els anys 1995 i 2003, d'aproximadament el 50%. Aquesta tendència, detectada també en altres comunitats autònomes (García-Fuster R et al, 2005), seria difícil d'interpretar sense suposar una important millora en la qualitat del procés assistencial, millora que es pot relacionar, en part, tant amb intervencions de gestió sanitària com amb la disseminació de resultats d'estudis com aquest.

Els resultats de la cirurgia coronària sense circulació extracorpòria

L'estudi dels resultats de la cirurgia sense CEC, més enllà de les implicacions que es discuteixen àmpliament en l'apartat de discussió de l'article, es pot comentar, en el context més ampli d'aquest treball, des de diferents punts de vista: 1) l'interès d'haver utilitzat una escala de risc per a fer una anàlisi

estratificada; 2) la il·lustració, un cop més, de com els estudis observacionals ens poden aportar un tipus d'informació diferent als assaigs clínics; i 3) els avantatges d'haver utilitzat un sistema d'ajust com l'índex de propensió.

Aquest estudi és un altre exemple de les múltiples possibles utilitats de l'ajust segons el risc. La possibilitat que ofereixen les escales de risc de resumir les característiques de risc més importants en una sola quantitat és un avantatge evident per a l'avaluació del resultat d'intervencions terapèutiques en estudis observacionals.

L'estratificació del risc segons una escala coneguda i àmpliament utilitzada en altres medis com és l'EuroSCORE és una novetat respecte d'altres estudis de comparació dels resultats entre la cirurgia sense CEC i la cirurgia amb CEC. El fet que la majoria d'assaigs clínics s'hagin realitzat en poblacions de risc preoperatori baix feia particularment interessant realitzar una anàlisi separada per als pacients de risc baix i de risc alt en un estudi observacional prospectiu. Aquesta anàlisi per separat ens ha permès observar una interacció interessant entre el perfil de risc i l'efecte de la cirurgia sense CEC sobre el resultat postoperatori. S'evidencia un avantatge important de la cirurgia sense CEC sobre diferents resultats en el postoperatori, i aquest avantatge es produeix sobretot a expenses de les diferències observades en el grup de pacients amb risc preoperatori baix.

Si bé el resultat global és assimilable al que s'observa en altres estudis observacionals, el fet que l'efecte sigui més important en el grup de pacients de baix risc és contradictori amb el resultat dels assaigs clínics. És poc habitual que les divergències entre els assaigs clínics i els estudis observacionals es produeixin en el sentit de detectar diferències més importants en els segons però, en aquest cas, la major part de l'evidència disponible prové d'estudis observacionals i els pocs assaigs clínics disponibles s'han realitzat sobre mostres relativament petites. Aquesta discrepància entre els assaigs clínics i els estudis observacionals pot ser deguda a diverses raons (Wijeysundera et al. 2005 i Ascione R et al, 2003): 1) la mortalitat en els assaigs clínics és menor en tractar-se de pacients de baix risc i la seva potència per detectar diferències és baixa; i 2) els criteris de selecció dels pacients sotmesos a cirurgia sense CEC, el tipus de tècnica utilitzada, el pas a CEC durant la intervenció, i l'experiència dels cirurgians son factors importants a tenir en compte en els estudis observacionals.

La selecció dels pacients es produeix sobre tot en funció de la preferència del cirurgià i de característiques anatòmiques (nombre de vasos afectats i tipus de vasos). De quina manera aquest biaix de selecció pot afectar el resultat en funció del risc preoperatori és difícil de deduir. En quan a les preferències dels cirurgians i equips quirúrgics, és plausible suposar que els que prefereixen aquesta tècnica, més innovadora, poden ser professionals, no necessàriament més experimentats, però sí millor qualificats i que han

desenvolupat l'habilitat necessària per aplicar-la de manera rutinària. Aquest fet pot relacionar-se d'altra banda amb el risc quirúrgic, perquè els cirurgians millor qualificats tendiran també a operar pacients de més risc, tant si la seva preferència és la cirurgia sense CEC com la cirurgia amb CEC. Tot això no es pot deduir directament de l'estudi però, alhora que permet interpretar els resultats observats, il·lustra: primer, la dificultat d'aplicar les tècniques d'aleatorització a una intervenció d'aquest tipus i; segon, la capacitat dels estudis observacionals realitzats de manera escurpulosa d'avaluar els resultats en condicions de pràctica clínica real. D'altra banda, és probable que aquest resultat reflecteixi una realitat assistencial no necessàriament limitada a l'àmbit de l'estudi sinó que, a la vista dels resultats obtinguts en altres estudis similars, pot ser molt probablement generalitzable a altres contextos.

Un aspecte a destacar d'aquest estudi és el fet d'haver utilitzat un mètode d'anàlisi que permet teòricament corregir el biaix de selecció en els estudis observacionals (D'Agostino RB, 1998). L'avantatge més important que té l'índex de propensió sobre altres tècniques d'anàlisi multivariant és la capacitat d'utilitzar models saturats (sense els requeriments de parsimònia que tenen els models logístics) per al seu càlcul. S'ha comprovat amb dades simulades que els índex de propensió proporcionen estimacions més precises de l'efecte que els models logístics quan és necessari, per fer un ajust correcte, incorporar en el model un nombre de covariables en una relació menor a 8 esdeveniments per covariable (Cepeda et al, 2003). Quan no és així, ambdues tècniques són igual d'eficients. Un altre avantatge important és la possibilitat d'explorar diferents resultats simultàniament sense haver de construir models diferents (amb diferents combinacions de factors), ja que en ser la probabilitat de tractament (l'índex de propensió) independent del resultat, podem utilitzar el mateix índex per a tots ells.

El repte més important consisteix doncs en estimar correctament i de manera no esbiaixada la probabilitat de tractament (l'índex de propensió). Si els factors de confusió importants no s'inclouen a l'anàlisi, igual que passa amb l'ajust directe mitjançant tècniques de regressió, l'estimació de l'efecte pot estar esbiaixat (Drake C et al, 1995). En el nostre estudi s'observava una tendència més gran a operar sense CEC pacients amb menys prevalença de factors de risc (diabetis i hipercolesterolèmia) i més antecedents de malaltia coronària, revascularització percutània, insuficiència cardíaca i disfunció neurològica, factors tots ells que poden també relacionar-se amb els resultats adversos però que, excepte la disfunció neurològica, no apareixen en les escales de risc utilitzades ni es relacionen de manera independent amb la mortalitat hospitalària en la mostra analitzada. Així doncs, aquests factors, que conjuntament poden afectar el resultat, no s'haurien utilitzat en un ajust del risc utilitzant les tècniques multivariants tradicionals però, en canvi, sí que s'han pogut tenir en compte juntament amb moltes altres en un anàlisi mitjançant l'índex de propensió.

Limitacions

Ja s'han descrit àmpliament totes les limitacions inherents a l'aplicació de qualsevol escala de risc i que per tant s'apliquen també a aquest estudi.

La limitació més important fa referència a l'exhaustivitat de la mostra. Encara que la mostra és exhaustiva pel que fa als centres privats, doncs són els únics a Catalunya que operen malalts procedents de la sanitat pública, inclou només tres dels cinc centres públics existents a Catalunya escollits en base a la seva accessibilitat a l'estudi. No es pot excloure que la inclusió de més centres diluís o reforçés les diferències detectades o que un període d'observació més llarg posés de manifest altres diferències ara menys patents.

El fet que les dues escales utilitzades no siguin específiques de cirurgia de revascularització coronària si no de cirurgia cardíaca en general (coronària i valvular) podria representar una limitació si la influència dels factors que són comuns en la predicció de la mortalitat fos diferent en els dos tipus de tècniques. Això podria influir en la calibració dels models. Parsonnet V et al (Parsonnet V et al, 1989), per exemple, van trobar que les característiques de validesa del seu model (desenvolupat a partir de les dades de cirurgia coronària i valvular) empitjoraven quan s'aplicava a procediments valvulars i coronaris per separat.

Per últim, en tractar-se d'un estudi observacional, encara que es va fer un control de qualitat estricte no es poden descartar errors sistemàtics en la recollida de dades.

Implicacions pràctiques i necessitats de recerca

Aquest treball, té tres tipus d'implicacions de caràcter ben diferent: 1) Actualitza la informació disponible sobre els resultats ajustats de la cirurgia coronària en les diferents institucions que la proveeixen en el marc de la sanitat pública catalana; 2) Afegeix nova informació sobre les estratègies més adequades per a l'ajust del risc de la cirurgia coronària en el nostre àmbit; i, 3) Contribueix a enriquir la noció de risc en les mentalitats de les persones implicades en la pràctica i la recerca clíniques.

Aquest treball ha permès constatar una millora important en els resultats de la cirurgia coronària en l'última dècada. Aquesta millora ha estat de la mateixa magnitud tant en centres de gestió pública com en centres de gestió privada. Com que els primers partien d'una situació pitjor encara s'observa una diferència marginal entre els dos tipus de gestió indicant que encara existeix un potencial de millora. L'avaluació periòdica d'aquests resultats en

col·laboració estreta amb els professionals de cada centre contribuiria a mantenir aquesta tendència de millora de la qualitat assistencial.

Arribats al punt en que ens trobem, quan ja fa algunes dècades que l'exercici de l'avaluació del risc ha esdevingut habitual arreu del món, i en el nostre país des de la iniciativa de l'AATRM, potser podríem aventurar-nos a establir les bases d'un sistema hipotètic d'avaluació del risc quirúrgic en l'àmbit català. L'estudi permet intuir quina seria l'estratègia més adequada per l'avaluació del risc en una iniciativa d'aquest tipus: l'ús d'un instrument d'ampli abast com l'EuroSCORE seria suficient per a la comparació de la mortalitat ajustada entre centres i l'ús d'un model local com el de l'AATRM ens oferiria la possibilitat d'una referència històrica.

No obstant, queden encara alguns temes per resoldre: 1) els estudis, com aquest, dissenyats per a un objectiu concret i realitzats de manera meticulosa, resulten costosos; i, 2) l'avaluació de la mortalitat hospitalària, encara que és fonamental, resulta insuficient quan ens interessa cada vegada més l'estat de salut del pacient i la seva qualitat de vida a més llarg termini.

Les dades per a l'ajust del risc poden obtenir-se de diverses fonts d'informació: dades de l'alta hospitalària en registres administratius, revisió retrospectiva de les històries clíniques o recollida prospectiva de dades primàries. La recollida prospectiva o retrospectiva (o una combinació de les dues) de dades primàries seguint un protocol sistematitzat és la manera més fiable de recollir la informació sobre el risc, però és també la que té majors requeriments de recursos econòmics i d'esforç personal. L'ús de bases de dades administratives seria el més adequat per a fer avaluacions periòdiques dels resultats en un sistema sanitari concret però té moltes limitacions: la informació sobre el risc no és prou completa (per exemple, no hi ha informació sobre la fracció d'ejecció o la prioritat de la intervenció), pot estar mal registrada i sovint no permet diferenciar les condicions cròniques de les que s'han produït durant l'ingrés, o entre esdeveniments que s'han produït abans o després de la intervenció quirúrgica. És poc probable doncs que les bases de dades administratives com el CMBD puguin substituir els estudis clínics rigorosos, però un anàlisi detallat de la seva validesa en comparació amb aquests ens podria ajudar a estimar la magnitud del problema i estudiar possibles formes de minimitzar-lo. Per exemple, és possible que es pogués obtenir una predicció del risc suficientment vàlida mitjançant la combinació de dades administratives i algunes, poques, dades clíniques recollides de forma prospectiva per l'equip quirúrgic durant el mateix procés assistencial.

En general hi ha poca informació sistemàtica sobre els resultats clínics a llarg termini. És habitual que s'utilitzi com a variable de resultat la mortalitat hospitalària, perquè és fàcil de recollir, no requereix un seguiment extrahospitalari i és altament discriminant com a indicador de mal resultat.

No obstant, seria de gran interès clínic conèixer quins son els determinants de mal resultat a més llarg termini i quin impacte té la intervenció sobre l'estat funcional i la qualitat de vida dels pacients. Per exemple, en l'estudi CIRCORCA, vam observar que la diferència de mortalitat entre centres de gestió pública i de gestió privada disminuïa al cap d'un any de la intervenció i que els predictors de resultat a curt i a llarg termini eren de naturalesa diferent. Mentre que la inestabilitat clínica (classificació del grau funcional i la indicació urgent o emergent) és el determinant més important de la mortalitat a curt termini, la presència de malalties cròniques, la revascularització incompleta i la percepció de salut física en el preoperatori augmenten en importància a l'hora de predir la mortalitat a més llarg termini. També la naturalesa d'aquests predictors és diferent segons el tipus de resultat que utilitzem. Per exemple, a l'estudi CIRCORCA, el fet de ser dona, tenir altres malalties concomitants i tenir pitjor salut autopercebuda (tant física com mental) en el preoperatori eren els predictors més importants de la percepció de pitjor salut física a l'any de la intervenció.

Per mirar de respondre a aquests dos problemes a la UESC s'està duent a terme actualment un projecte de recerca finançat pel FIS que representa una continuació de l'estudi ARCA i ens permetrà: 1) Analitzar la validesa de les bases de dades administratives basades en el Conjunt Mínim Bàsic de Dades (CMBD) per a la predicció del risc quirúrgic en comparació amb una base de dades clínica i l'ús d'escala de risc específiques. 2) Analitzar el resultat de la cirurgia coronària a més llarg termini (5 anys) en termes de resultat clínic i qualitat de vida relacionada amb la salut.

Per a l'avaluació de la qualitat assistencial, l'aplicació d'instruments de predicció del risc és fonamental. No sempre ha estat ben rebut, per part dels professionals implicats, l'anàlisi i la difusió dels resultats de la cirurgia cardíaca, sobretot per la manca de diàleg entre clínics, metodòlegs i gestors (Hannan EL et al, 1997). No obstant, sigui quin sigui el sistema escollit per a l'ajust del risc, l'exercici de comparar el resultat ajustat entre centres, equips quirúrgics o sistemes de gestió, pot ser útil per millorar la qualitat assistencial quan els resultats retornen als mateixos proveïdors per induir i facilitar canvis en el procés assistencial que repercuteixin en el resultat (Weintraub WS, 2006). Hi ha algunes experiències documentades sobre aquest *feed-back* positiu (O'Connor GT et al, 1996 i Hannan EL, 1994) encara que no està clar que la difusió pública dels resultats per centres i equips quirúrgics hagi contribuït significativament i de manera directa a la millora de la qualitat assistencial (Ghali WWA et al, 1997). En la nostra experiència, estem convençuts que l'esforç realitzat, la difusió dels resultats en diferents cercles (tant clínics com metodològics) i l'intercanvi amb els investigadors dels diferents centres participants, ha contribuït a fomentar el diàleg i l'autocrítica dels equips quirúrgics i dels diferents professionals implicats.

Finalment, val la pena comentar que l'anàlisi del risc igual com l'anàlisi d'altres mesures d'interès clínic (per exemple, la qualitat de vida relacionada amb la salut, les escales funcionals o les escales de gravetat de determinades patologies), en les que s'utilitzen sistemes quantitius per a la mesura de conceptes més o menys abstractes, té generalment un impacte qualitatiu interessant: contribueix a enriquir la noció intuïtiva del concepte analitzat i proporciona un llenguatge comú per a la difusió, discussió i interpretació crítica de la mesura. Aquests conceptes més o menys abstractes s'anomenen constructes perquè, tot i que intuïtivament tenen una entitat pròpia, no es poden mesurar directament si no a partir de la mesura de diferents aspectes que s'hi relacionen. El "risc quirúrgic", malgrat tractar-se d'un probabilitat, també té les característiques de constructe pel fet de ser un estat susceptible de ser quantificat mitjançant la mesura de diferents indicadors que s'hi relacionen. Així doncs, l'estudi dels determinants del risc i l'ús d'instruments específics per a la mesura del risc quirúrgic, encara que no puguin aplicar-se com a criteri únic per a prendre decisions clíniques, té un impacte qualitatiu sobre la mentalitat dels professionals de la salut que pot influir significativament en la seva pràctica diària.

CONCLUSIONS

Sobre l'avaluació dels resultats de la cirurgia de revascularització coronària a Catalunya:

- 1) La mortalitat hospitalària d'un primer implant aortocoronari en els pacients de la sanitat pública catalana és, tant en centres públics com privats, equivalent o inferior a l'esperada segons dos instruments d'ajust del risc.
- 2) No obstant, l'ajust segons les característiques basals dels pacients suggereix l'existència d'una diferència favorable als centres de gestió privada i compatible amb diferències en el procés assistencial.
- 3) La disminució de la mortalitat respecte estudis previs permet constatar una tendència general de millora del procés assistencial en la cirurgia coronària a Catalunya en els últims anys.

Sobre les escales de predicció del risc:

- 1) El contingut de les escales més enllà de certs trets comuns a totes elles (l'edat, la presència de malalties concomitants i algun indicador de severitat) té poc impacte en la validesa de les escales per al seu ús com a instrument d'ajust del risc o com a criteri general de qualitat.
- 2) La diferència de mortalitat entre la població de referència i la població estudiada té un impacte fonamental en la calibració de l'escala que cal interpretar en funció de l'aplicació que se'n faci.
- 3) Les escales d'àmbit local i d'àmbit general tenen propietats complementàries per a l'avaluació del risc de la cirurgia coronària.

Sobre l'efectivitat de la cirurgia coronària sense circulació extracorpòria:

- 1) En condicions de pràctica clínica real, la cirurgia coronària sense circulació extracorpòria pot associar-se a menys esdeveniments desfavorables que la cirurgia amb circulació extracorpòria
- 2) En contrast amb estudis previs, aquest benefici s'observa en els pacients de baix risc preoperatori

A part de les conclusions específiques de cada un dels estudis, de la visió conjunta de tot el treball realitzat se'n deriven les següents conclusions generals:

- 1) L'ajust del risc permet observar canvis en la pràctica clínica, facilita la interpretació crítica dels resultats i és útil per a l'estratificació del risc en estudis d'efectivitat.
- 2) L'ajust del risc ha enriquit la nostra noció del risc com a entitat concreta i objectiva susceptible d'anàlisi rigorós.

BIBLIOGRAFIA

Ascione R, Caputo M, Angelini GD. Off-pump coronary artery bypass grafting: not a flash in the pan. *Ann Thorac Surg* 2003;75:306-13

Babapulle MN, Joseph L, Bélisle P, Brophy JM, Eisenberg MJ. A hierarchical Bayesian meta-analysis of randomised clinical trials of drug-eluting stents. *Lancet* 2004;364:583-91

BARI Trial Investigators. Five-year clinical and functional outcome comparing bypass surgery and angioplasty in patients with multivessel coronary disease. A multicenter randomized trial (BARI trial). *JAMA* 1997;277:715-21

BARI Trial Investigators. Seven-year outcome in Bypass Angioplasty Revascularization Investigation (BARI) by treatment and diabetic status. *J Am Coll Cardiol* 2000;35:1122-9

Cepeda MS, Boston R, Farrar JT, Storm BL. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *Am J Epidemiol* 2003;158:280-287

Cheng DC, Bainbridge D, Martin JE, Novick RJ. Does off-pump coronary artery bypass reduce mortality, morbidity, and resource utilization when compared with conventional coronary artery bypass? A meta-analysis of randomized trials. *Anesthesiology* 2005;102:188-203

Cortina-Romero JM. Scores de gravedad y complejidad en cirugía cardíaca. Usos y limitaciones. *Rev Esp Cardiol* 2005; 58: 473-476

D'Agostino RB, Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* 1998;17:2265-2281

Detre KM, Guo P, Holubkov R, Califf RM, Sopko G, Bach R, Brooks MM, Bourassa MG, Shemin RJ, Rosen AD, Krone RJ, Frye RL, Feit F. Coronary revascularization in diabetic patients: a comparison of the randomized and observational components of the Aypass Angioplasty Revascularization Investigation (BARI). *Circulation* 1999;99:633-40

Drake C, Fisher L. Prognostic models and the propensity score. *Int J Epidemiol* 1995;24:183-187

Edwards FH, Albus RA, Zajtchuk R, Graeber GM, Barry MJ, Rumisek JD, Arishita G. Use of a Bayesian statistical model for risk assessment in coronary artery surgery. *Ann Thorac Surg* 1988;45:437-440

Edwards FH, Clark RE, Schwartz M. Coronary artery bypass grafting: the society of Thoracic Surgeons national database experience. *Ann Thorac Surg* 1994;57:12-9

García-Fuster R, Montero JA, Gil O, Hornero F, Cánovas S, Bueno M, et al. Tendencias en cirugía coronaria: cambios en el perfil del paciente quirúrgico. *Rev Esp Cardiol* 2005;58:512-522

Gersh BJ, Frye RL. Methods of coronary revascularization – Things may not be as they seem. *N Engl J Med* 2005;352:2235-7

Ghali WWA, Ash AS, Hall RE, Moskowitz MA. Statewide quality improvement initiatives and mortality after cardiac surgery. *JAMA* 1997;277:379-382

González-Santos JM. ¿Gestión pública o gestión privada? Implicaciones en los resultados de la cirugía coronaria. *Rev Esp Cardiol* 2006;59:414-7

Hannan EL, Kilburn H, O'Donnell JF, Lukacik G, Shields EP. Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates. *JAMA* 1990;264:2768-74

Hannan EL, Kilburn H, Racz M, Shields E, Chassin MR. Improving the outcomes of coronary artery bypass surgery in New York State. *JAMA* 1994;271:761-766

Hannan EL, Racz MJ, Wallford G, et al. Long-term outcomes of coronary-artery bypass grafting versus stent implantation. *N Engl J Med* 2005;352:2174-83

Hannan EL, Stone CC, Biddle TL, DeBuono BA. Public release of cardiac surgery outcomes data in New York: what do New York state cardiologists think of it? *Am Heart J* 1997;134:1120-8

Hannan EL, Wu C. Assessing quality and outcomes for percutaneous coronary intervention: choosing statistical models, outcomes, time periods, and patient populations. *Am Heart J*. 2003;145:571-4

Hannan EL, Wu C, Bennett EV, Carlson RE, Culliford AT, Gold JP, Higgins RSD, Isom OW, Smith CR, Jones RH. Risk stratification of in-hospital mortality for coronary artery bypass graft surgery. *J Am Coll Cardiol* 2006;47:661-8

Herlitz J, Wiklund I, Caidahl K, Karlson BW, Sjöland H, Hartford M, Haglid M, Karlsson T. Determinants of an impaired quality of life five years after coronary artery bypass surgery. *Heart* 1999;81:342-346

Higgins TL, Estafanous FG, Loop FD, Beck GJ, Blum JM, Paranandi L. Stratification of morbidity and mortality outcome by preoperative risk factors in coronary artery bypass patients: a clinical severity score. *JAMA* 1992;267:2344-8

Hlatky MA, Boothroyd DB, Melsop KA, Brooks MM, Mark DB, Pitt B, Reeder GS, Rogers WJ, Ryan TJ, Whithlow PL, Wiens RD. Medical costs and quality of life 10 to 12 years after randomization to angioplasty or bypass surgery for multivessel coronary artery disease. *Circulation* 2004;110:1960-6

Iezzoni LI, Ash AS, Ahwartz M, Landon BE, Mackiernan YD. Predicting in-hospital deaths from coronary artery bypass graft surgery: do different severity measures give different predictions? *Med Care* 1998;36:28-39

Iezzoni LI, Ash AS, Shwartz M, Daley J, Hughes JS, Mackiernan YD. Predicting who dies depends on how severity is measured: implications for evaluating patient outcomes. *Ann Intern Med* 1995;123:763-70

Iezzoni LI. Using risk-adjusted outcomes to assess clinical practice: an overview of issues pertaining to risk adjustment. *Ann Thorac Surg* 1994;58:1822-6

Ivanov J, Tu JV, Naylor CD. Ready-made, recalibrated, or remodeled? Issues in the use of risk indexes for assessing mortality after coronary artery bypass graft surgery. *Circulation* 1999;99:2098-2104

Krumholz HM. Mathematical models and the assessment of performance in cardiology. *Circulation* 1999;99:2067-2069

Landon B, Iezzoni LI, Ash AS, et al. Judging hospitals by severity-adjusted mortality rates: the case of CABG surgery. *Inquiry* 1996;33:155-66

Lemos PA, Serruys PW, van Domburg RT, Saia F, Arampatzis CA, Hoye A, Degertekin M, Tenabe K, Daemen J, Liu TTK, McFadden E, Sianos G, Hofma SH, Smits PC, van der Giessen WJ, de Feyer PJ. Unrestricted utilization of sirolimus-eluting stents compared with conventional bare stent implantation in the "real world". *Circulation* 2004;109:190-195

Nashef SAM, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R, and the EuroSCORE study group. European system for Cardiac Operative Risk Evaluation (EuroSCORE). *Eur J Cardiothorac Surg* 1999;16:9-13

Nathoe HM, van Dijk D, Jansen EW, Suyker WJ, Diephuis JC, van Boven WJ, de la Riviere AB, Borst C, Kalkman CJ, Grobbee DE, Buskens E, de Jaegere

PP. A comparison of on-pump and off-pump coronary bypass surgery in low-risk patients. *N Engl J Med* 2003;348:394-402

O'Connor GT, Plume SK, Olmstead EM, Morton JR, Maloney CT, Nugent WC, Hernandez F Jr, Clough R, Leavitt BJ, Coffin LH, Marrin CA, Wennberg D, Birkmeyer JD, Charlesworth DC, Malenka DJ, Quinton HB, Kasper JF. A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. The Northern New England Cardiovascular Disease Study Group. *JAMA* 1996;275:841-6

Parsonnet V, Dean D, Bernstein AD. A method for uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. *Circulation* 1989;79 Suppl 1:13-12

Permanyer-Miraldà G, Brotons C, Cascant O, Moral I, Alonso J, Soler Soler J. Valoración de la calidad de vida relacionada con la salud a los dos años de la cirugía coronaria. *Med Clin (Barc)*, 1997;108:446-451

Permanyer-Miraldà G, Brotons C, Ribera A, Alonso J, Cascant P, Moral I y grupo de investigadores del estudio CIRCORCA. Resultados después de cirugía coronaria: determinantes de calidad de vida relacionada con la salud postoperatoria. *Rev Esp Cardiol* 2001; 54: 607-616

Permanyer-Miraldà G, Brotons C, Ribera A, Cascant P, Pons JMV, Alonso J et al. Desigual perfil clínico, calidad de vida y mortalidad hospitalaria en pacientes operados de injerto aortocoronario en centros públicos y privados de Cataluña. *Rev Esp Cardiol* 1998;51:806-815

Permanyer-Miraldà G, Brotons C, Ribera A, Moral I, Cascant P, Alonso J, Martínez C, por el estudio QUALISTENT. Resultados clínicos y de calidad de vida de los pacientes tratados con angioplastia coronaria con balón o stent. Estudio multicéntrico prospectivo. *Rev Esp Cardiol* 2001; 54: 597-606

Peterson ED, DeLong ER, Muhlbauer LH, Rosen AB, Buell HE, Kiefe CI, Kresovik TF. Challenges in comparing risk-adjusted bypass surgery mortality results. *J Am Coll Cardiol* 2000;36:2174-84

Pitkänen O, Niskanen M, Rehnberg S, Hippeläinen M, Hynynen M. Intra-institutional prediction of outcome after cardiac surgery: comparison between a locally derived model and the EuroSCORE. *Eur J Cardio-thorac Surg* 2000;18:703-10

Pons JMV, Borrás JM, Espinas JA, Moreno V, Cardona M, Granados A. Subjective versus statistical model assessment of mortality risk in open heart surgical procedures. *Ann Thorac Surg* 1999; 67: 635-40

Pons JMV, Espinas JA, Borrás JM, Moreno V, Martín I, Granados A. Cardiac surgical mortality. Comparison among different additive risk-scoring models in a multicenter sample. *Arch Surg* 1998;133:1053-7

Pons JMV, Granados A, Espinas JA, Borrás JM, Martín I, Moreno V. Assessing open heart surgery mortality in Catalonia (Spain) through a predictive risk model. *Eur J Cardio-thoracic Surg* 1997; 11: 415-423

Pons JMV, Moreno V, Borrás J, Espinàs J, Amazan C, Granados A. Open Heart surgery in public and private practice. *J Health Serv Res Policy* 1999;4:73-8

Rihal CS, Raco DL, Gersh BJ, Yusuf S. Indications for coronary artery bypass graft surgery and percutaneous coronary intervention in chronic stable angina: review of the evidence and methodological considerations. *Circulation* 2003;108:2439-45

Roques F, Nashef P, Michel P, Pinna Pintor P, David M, Baudet E for the EuroSCORE study group. Does EuroSCORE work in individual European countries? *Eur J Cardiothorac Surg* 2000;18:27-30

Rumsfeld JS, MaWinney S, McCarthy M, et al. Health related quality of life as a predictor of mortality following coronary artery bypass graft surgery. *JAMA* 1999;281:1298-303

Shroyer ALW, Coombs LP, Peterson ED, Eiken MC, DeLong ER, Chen A, Ferguson TB, Grover FL, Edwards FH. The Society of Thoracic Surgeons: 30-day operative mortality and morbidity risk models. *Ann Thorac Surg* 2003;75:1856-65

Steyerberg EW, Eijkemans MJC, Boersma E, Habbema JDF. Equally valid models gave divergent predictions form mortality in acute myocardial infarction patients in a comparison of logical regression models. *J Clin Epidemiol* 2005;58:383-90

Stürmer T, Schneeweiss S, Brookhart MA, Rothman KJ, Avorn J, Glynn RJ. Analytic strategies to adjust confounding using exposure propensity scores and disease risk scores: nonsteroidal antiinflammatory drugs and short-term mortality in the elderly. *Am J Epidemiol* 2005;161:891-898

Tu JV, Jaglal SB, Naylor CD. Multicenter validation of a risk index for mortality, intensive care unit stay, and overall hospital length of stay after cardiac surgery. Steering Committee of the Provincial Adult Cardiac Care Network of Ontario. *Circulation*. 1995;91:677-84

Tu JV, Sykora K, Naylor CD. Assessing outcomes of CABG: How many risk factors are enough? *JACC* 1997; 30: 1317-23

Weintraub WS. Evaluating the risk of coronary surgery and percutaneous coronary intervention. *J Am Coll Cardiol* 2006;47:669-671

Wijeyesundera DN, Beattie WS, Djaiani G, Rao V, Borger MA, Karkouti K, Cusimano RJ. Off-pump coronary artery surgery for reducing mortality and morbidity: meta-analysis of randomized and observational studies. *J Am Coll Cardiol* 2005;46:872-882

Yusuf S, Zucker D, Peduzzi P, Fisher LD, Takaro T, Kennedy JW, et al. Effect of coronary artery bypass graft surgery on survival: overview of 10-year results from randomised trials by the Coronary Artery Bypass Graft Surgery Trialists Collaboration. *Lancet* 1974;334:563-70

ANNEX I. Estudi CIRCORCA

Permanyer-Miralda G, Brotons C, Ribera A, Cascant P, Pons JMV, Alonso J et al. Desigual perfil clínico, calidad de vida y mortalidad hospitalaria en pacientes operados de injerto aortocoronario en centros públicos y privados de Cataluña. Rev Esp Cardiol 1998;51:806-815

Permanyer-Miralda G, Brotons C, Ribera A, Moral I, Cascant P, Alonso J, Martínez C, por el estudio QUALISTENT. Resultados clínicos y de calidad de vida de los pacientes tratados con angioplastia coronaria con balón o stent. Estudio multicéntrico prospectivo. Rev Esp Cardiol 2001; 54: 597-606

Permanyer-Miralda G, Brotons C, Ribera A, Cascant P, Pons JMV, Alonso J, Mallol Kirchner A, Barthe Carrera J, Martínez Useros C, Cardona Burrull M, Delago Ramis L, Romero Ferrer B (en representación del estudio CIRCORCA)

Desigual perfil clínico, calidad de vida y mortalidad hospitalaria en pacientes operados de injerto aortocoronario en centros públicos y privados de Cataluña

Revista Española de Cardiología 1998; 51(10):806-815

Permanyer-Miralda G, Brotons C,. Ribera A, Alonso J, Cascant P, Moral Perez I (en representación del estudio CIRCORCA)
Resultados después de cirugía coronaria: determinantes de calidad de vida relacionada con la salud postoperatoria
Revista Española de Cardiología 2001; 54(5):606-616

ANNEX 2. Formulari de recollida de dades i definicions de les variables d'estudi

**Estudi multicèntric per a l'avaluació dels
resultats de la cirurgia coronària ajustats
al risc a Catalunya
(Estudi ARCA)**

Unitat d'Epidemiologia Clínica
Servei de Cardiologia
Hospital Universitari Vall d'Hebron

Questionari de recollida de dades
(octubre 2001 – setembre 2003)

Verificació dels criteris d'inclusió:

- **Primera intervenció de derivació aorto-coronària**
- **No hi ha altres operacions cardíaques associades**
- **Finançament públic de la intervenció**

Dades d'identificació del malalt

1. N° d'ordre	<input type="text"/>
2. Centre participant	<input type="text"/>
(1)Hospital Vall d'Hebron (2)Hospital Germans Trias i Pujol (3)Hospital de Sant Pau (4)Hospital General de Catalunya (5)Centre Cardiovascular Sant Jordi	
3. Número d'història clínica	<input type="text"/>
4. Nom i cognoms	<input type="text"/>
5. Telèfons	<input type="text"/> <input type="text"/>

Dades demogràfiques

6. Sexe	<input type="text"/>
(1)home (2)dona	
7. Data de naixement	<input type="text"/>
8. Edat	<input type="text"/>
9. Situació laboral	<input type="text"/>
(1)Actiu (2)Jubilat (3)Aturat (4)Baixa o invalidesa (5)Mestressa de casa (6)Altres	
10. Nivell d'estudis	<input type="text"/>
(1)Sense estudis (2)Estudis primaris incomplets (3)Estudis primaris o EGB fins a 5è (4)Graduat escolar, EGB fins a 8è, batxillerat elemental o similar (5)Batxillerat superior, BUP, FP o similar (6)Estudis universitaris	

11. Es visita pel cardiòleg del seu ambulatori o a les consultes	<input type="text"/>
(1)Regularment (2)De vegades (3)No (9)No consta	
12. Es visita per un cardiòleg privat?	<input type="text"/>
(1)Regularment (2)De vegades (3)No (9)No consta	
13. Desde quan es visita pel cardiòleg?	<input type="text"/>
14. Remès des d'un hospital privat	<input type="text"/>
(1)Sí (2)No (9)No consta	

Factors de risc cardiovascular

15. Hipertensió arterial	<input type="text"/>
(1)Sí (2)No (9)No consta	
16. Hipercolesterolèmia	<input type="text"/>
(1)Sí (2)No (9)No consta	
17. Diabetis	<input type="text"/>
(1)Sí (2)No (9)No consta	
18. Fumador	<input type="text"/>
(1)Sí, fumador actual (2)No fumador (3)Exfumador (> 1 any)	

Antecedents (previs a l'ingrés actual)

19. Antecedents clínics de cardiopatia isquèmica	<input type="text"/>
(1)Sí (2)No	
20. Antecedents de revascularització	<input type="text"/>
(1)Sí (2)No	
21. Antecedents d'insuficiència cardíaca	<input type="text"/>
(1)Sí (2)No	
22. Antecedents d'AVC	<input type="text"/>
(1)Sí (2)No	

Determinants del risc en el moment de la indicació quirúrgica

23. DA proximal

(1)Sí (2)No (9)No consta

24. Nombre de vasos

25. Data de la indicació quirúrgica

26. Data del debut de la malaltia coronària aguda que motiva la indicació

(només per IAM i angina inestable)

27. Creatinina ($\geq 1,5$ mg/dl)

(1)Sí (2)No (9)No consta

28. Disfunció neurològica

(1)Sí (2)No

29. Malaltia hepàtica

(1)Sí (2)No

30. Malaltia pulmonar crònica

(1)Sí (2)No

31. Aneurisma ventricular esquerre

(1)Sí (2)No

32. Claudicació intermitent

(1)Sí (2)No

33. Oclusió o estenosi de $>50\%$ d'una artèria caròtida

(1)Sí (2)No (9)No consta

34. Intervenció (prèvia o programada) de l'aorta abdominal, caròtides o extremitats

(1)Sí (2)No

35. Grau funcional CCS

0, I, II, III, IV

36. Hipertensió pulmonar sistòlica (>60 mmHg)

(1)Sí (2)No (9)No consta

Disfunció ventricular

37. esquerra

(1)normal($>50\%$)

(2)lleugera-moderadament deprimida(30-50%)

(3)severament deprimida($<30\%$)

(9)no consta

38. Infart de miocardi

(1)Sí (2)No

39. Data de l'últim infart

Esdeveniments durant el temps d'espera

40. Infart

(1)Sí (2)No (9)No consta

41. Data de l'últim infart

42. Ingress per angina inestable*

(1)Sí (2)No (9)No consta

43. Nombre d'ingressos per malaltia coronària*

44. Nombre d'atencions a urgències per malaltia coronària*

45. Mort

(1)Sí (2)No

46. Data de la defunció

47. Causa cardiovascular

(1)Sí (2)No (9)No consta

48. Refús de la operació

(1)Per part del malalt

(2)Per part del cirurgià (3)No

49. S'opera en un altre centre

(1)Sí (2)No (9)No consta

50. Demora en la intervenció

(1)Sí (2)No (9)No consta

51. Motiu de la demora

*Només procedeix en malalts electius

Motiu de l'ingrés actual

Data inicial d'ingrés (en cas de trasllat des d'un altre centre)	<input type="text"/>
52.	
Data de l'ingrés 53. actual	<input type="text"/>
54. Cirurgia programada (1)Sí (2)No (9)No consta	<input type="text"/>
55. Angina inestable (1)Sí (2)No (9)No consta	<input type="text"/>
56. Infart de miocardi (1)Sí (2)No (9)No consta	<input type="text"/>
57. Insuficiència cardíaca (1)Sí (2)No (9)No consta	<input type="text"/>

Determinants del risc en el moment de la intervenció

58. Data de la intervenció	<input type="text"/>
59. Indicació (1)Electiva (2)Urgent (3)Emergent	<input type="text"/>
Última crisi d'angina de repòs en les 72 hores prèvies 60.	<input type="text"/>
(1)Sí (2)No	
61. Creatinina (≥ 1.5 mg/dl) (1)Sí (2)No (9)No consta	<input type="text"/>
62. Disfunció neurològica (1)Sí (2)No	<input type="text"/>
63. Malaltia hepàtica (1)Sí (2)No	<input type="text"/>
64. Malaltia pulmonar crònica (1)Sí (2)No	<input type="text"/>
Aneurisma ventricular esquerre 65.	<input type="text"/>
(1)Sí (2)No	
66. Claudicació intermitent (1)Sí (2)No	<input type="text"/>
Oclusió o estenosi de $>50\%$ d'una artèria caròtida 67.	<input type="text"/>
(1)Sí (2)No (9)No consta	
Intervenció (prèvia o programada) de l'aorta abdominal, caròtides o extremitats 68.	<input type="text"/>
(1)Sí (2)No	

69. Grau funcional CCS 0, I, II, III, IV	<input type="text"/>
Hipertensió pulmonar sistòlica (>60mmHg) 70.	<input type="text"/>
(1)Sí (2)No (9)No consta	
Disfunció ventricular esquerra 71.	<input type="text"/>
(1)Normal ($>50\%$) (2)Moderadament deprimida(30-50%) (3)Severament deprimida($<30\%$) (9)No consta	
72. Xoc cardiogènic	<input type="text"/>
(1)Sí (2)No	
Angina inestable: Nitrats i.v. fins a l'entrada a quiròfan 73.	<input type="text"/>
(1)Sí (2)No	
Taquicardia ventricular o fibril·lació 74.	<input type="text"/>
(1)Sí (2)No	
Massatge cardíac preoperatori 75.	<input type="text"/>
(1)Sí (2)No	
Intubació abans de l'entrada a quiròfan 76.	<input type="text"/>
(1)Sí (2)No	
Suport ionotròpic intravenós (preop) 77.	<input type="text"/>
(1)Sí (2)No	
Contrapulsació intraaòrtica 78.	<input type="text"/>
(1)Sí (2)No	
Insuf renal aguda preop.: Anúria o oligonúria < 10ml/h. 79.	<input type="text"/>
(1)Sí (2)No	
80. Ruptura septal postinfart (1)Sí (2)No	<input type="text"/>
Nombre d'empelts:	
81. Mamària	<input type="text"/>
82. Radial	<input type="text"/>
83. Safena	<input type="text"/>
Cirurgia sense circulació extracorpòria 84.	<input type="text"/>
(1)Sí (2)No	

Resultats. Postoperatori immediat

85. Sagnament

(1)Sí (2)No

86. Infart peroperatori

(1)Sí (2)No

87. Reintervenció

(1)Sí (2)No

Complicacions infeccioses

(ferida esternal,

88. respiratòries, nosocomials)

(1)Sí (2)No

89. Quina

Altres complicacions greus

90. cardiovasculars

(1)Sí (2)No

91. Quines

Altres complicacions greus

92. no cardiovasculars

(1)Sí (2)No

93. Quines

94. Mort hospitalària

(1)Sí (2)No

95. Data d'alta o defunció

Resultats. A l'any de la intervenció

96. Data de l'enquesta

97. Angina

(1)Sí (2)No

98. Grau funcional CCS

0, I, II, III, IV

99. Ingrés per causa cardiovascular

(1)Sí (2)No

100. ACTP

(1)Sí (2)No

101. Cirurgia

(1)Sí (2)No

102. Mort

(1)Sí (2)No

103. Data de la defunció

104. Causa cardiovascular

(1)Sí (2)No (9)No consta

**Estudi multicèntric per a l'avaluació dels resultats de la cirurgia
coronària ajustats al risc a Catalunya
(Estudi ARCA)**

Unitat d'Epidemiologia Clínica. Servei de Cardiologia.
Hospital Universitari Vall d'Hebron

Manual de procediments
(octubre 2001 – setembre 2003)

Normes generals

Aquest és un manual que pretén ser un ajut per omplir les dades del qüestionari ARCA. Hem procurat que el nombre de variables sigui el mínim possible per tal de que la recollida de dades sigui el més acurada possible. El qüestionari consta de tres parts segons el moment en que fan referència les dades: el moment de la programació, el moment de la operació i el seguiment.

És recomanable, sempre que sigui possible, omplir les dades d'identificació del malalt i demogràfiques (1-14), els factors de risc cardiovascular (15-18), els antecedents (19-22) i els determinants del risc en el moment de la indicació (23-39), quan el malalt es programa per a cirurgia. En cas que el malalt es programi de forma electiva, arxivarem el qüestionari i el recuperarem per omplir les dades d'esdeveniments durant el temps d'espera (40-51), el motiu de l'ingrés (52-56), els determinants del risc en el moment de la intervenció (57-83) i els resultats del postoperatori immediat (84-92) quan el malalt ingressi per a operar-se. En operacions urgents i emergents, la indicació i la cirurgia estan dins el mateix ingrés, de manera que és poc probable que es produeixin canvis en el risc quirúrgic, així no obstant omplirem les variables 23-39, encara que algunes siguin redundants amb els determinants del risc en el moment de la intervenció. Tornarem a recuperar el qüestionari un any després de la cirurgia, moment en el que caldrà determinar els resultats a l'any de la intervenció (93-101).

És important no perdre cap malalt que compleixi els criteris d'inclusió i intentar recollir el màxim d'informació de la forma més acurada possible, mitjançant la revisió d'històries clíniques, l'entrevista amb el malalt, o la consulta de possibles registres dels serveis de cirurgia que cada hospital pot tenir.

La persona responsable del treball de camp marcarà amb una X la resposta que correspongui en les preguntes tancades (p.ex. el sexe o la situació laboral) i a continuació apuntarà el número que correspongui en la casella de codificació. En les preguntes obertes (p.ex. les dates, el telèfon o el nom del malalt) apuntarà el que correspongui en la casella de codificació. En el cas de les dates, s'apuntarà una data aproximada o molt probable en el cas de no saber la data exacta, o com a mínim el mes i l'any. No es pot deixar en blanc cap de les variables, en cas de no trobar la informació, es codificarà amb un 9 (no consta), o es consultarà amb el centre coordinador des d'on farem el possible per aclarir qualsevol dubte o per corregir mancances o incongruències en el qüestionari. En aquest sentit, seria important dedicar una setmana a fer un estudi pilot, recollint la informació en uns quants malalts abans de començar l'estudi, per tal d'aclarir dubtes i detectar possibles problemes.

Verificació dels criteris d'inclusió:

Abans de registrar l'entrada a l'estudi d'un malalt, caldrà verificar si es compleixen els criteris d'inclusió i es marcaran amb un √:

- Primera intervenció de derivació aorto-coronària
- No hi ha altres operacions cardíaques associades
- Finançament públic de la intervenció: entraran a l'estudi tots aquells malalts en que la cirurgia es dugui a terme amb finançament públic. Entren en aquesta categoria les intervencions finançades pel PAMEM (Mútua municipal). En les preguntes 11 a 14 quedarà clar si el malalt és atès a més a més a la sanitat privada.

Dades d'identificació del malalt:

1. Centre participant
2. Número d'identificació del malalt: assignació consecutiva de números enters als pacients segons l'ordre en que són incorporats al protocol d'estudi, és a dir, així que son programats per a la cirurgia. Es començarà desde 1 en cada hospital.
3. Número d'història clínica
4. Nom i cognoms: en majúscules
5. Teléfon: posar el prefix. Indicar si és el teléfon d'un familiar o persona de contacte. És important tenir una manera de localitzar el malalt (posar més d'un teléfon si és necessari), per tal de poder recollir la informació en el seguiment.

Dades demogràfiques

6. Sexe
7. Data de naixement: data completa (dia/mes/any) (p.ex. 26/02/1950)
8. Edat: en el moment de la programació
9. Nivell d'estudis: els diplomats o les carreres intermitges es consideraran estudis universitaris
10. Situació laboral
11. Es visita pel cardiòleg de l'ambulatori
12. Es visita per un cardiòleg privat
13. Desde quan es visita per un cardiòleg, ja sigui privat o de l'ambulatori. Posar la data aproximada p.ex si es fa la pregunta al malalt i aquest ens diu que fa 3 anys: posarem el dia 15, del mes actual, de fa tres anys.
14. Indicarem aquí si el malalt ha estat traslladat desde un hospital privat.

Factors de risc cardiovascular

15. Hipertensió arterial: si consta a la història clínica o si en alguna ocasió s'ha notificat al malalt la seva condició d'hipertens o si està en tractament antihipertensiu
16. Hipercolesterolèmia: si consta a la història clínica o si en alguna ocasió s'ha notificat al malalt la seva condició d'hipercolesterolèmic o si està en tractament hipocolesterolemiant.

17. Diabetis: si consta a la història o si en alguna ocasió s'ha notificat al pacient la seva condició de diabètic o si està en tractament amb insulina o antidiabètics orals.

18. Fumador: es considera fumador a partir d'una cigarreta al dia (inclosos puros i pipa). Considerarem exfumador al menys un any sense fumar o si consta a la història com a exfumador.

Antecedents (previs a l'ingrés actual)

Indicarem si hi ha o no antecedents de malaltia cardiovascular: cardiopatia isquèmica, revascularització (ACTP o STENT), insuficiència cardíaca o AVC abans de l'ingrés que motiva la cirurgia. L'antecedent específic d'infart apareix més endavant, a la casella 38, perquè l'infart intervé en les escales de risc quirúrgic que volem valorar i és important registrar la data el més exactament possible, tant si és un antecedent, com si és la síndrome que motiva l'ingrés, com si es produeix durant l'ingrés abans de la cirurgia. .

Determinants del risc en el moment de la indicació quirúrgica

Avaluar el risc quirúrgic, tant en el moment de la indicació com de la cirurgia és l'objectiu més important de l'estudi, per tant, s'ha de tenir especial cura de recollir la informació sobre els determinants del risc. Hi ha alguns determinants que si no apareixen a la història podrem deduir fàcilment que no existeixen, com p.ex. disfunció neurològica, malaltia hepàtica, EPOC, etc., per tant, hem posat la opció (9) No consta, només en aquelles variables que requereixen una exploració per a confirmar-les, com és el cas de la creatinina, la oclusió de la caròtida i la hipertensió pulmonar.

23. DA proximal.

24. Nombre de vasos amb estenosi significativa (>50%)

25. Data de la indicació quirúrgica: data completa (dia/mes/any) (p.ex. 18/05/2001) en la que es presenta el malalt en la sessió mèdico-quirúrgica

26. Data del debut de la síndrome coronària aguda que motiva la indicació. Omplir només en cas d'indicació per angina inestable o IAM

27. Creatinina ($\geq 1,5$ mg/dl)

28. Disfunció neurològica: Afectació severa de la deambulació, o del llenguatge, o del nivell cognitiu, o de l'activitat diària

29. Malaltia hepàtica: Diagnòstic clínic confirmat o histològic d'hepatitis crònica o cirrosi hepàtica.

30. Malaltia pulmonar crònica: ús habitual de broncodilatadors o esteroides o necessitat d'ingressos

31. Aneurisma ventricular esquerre: operat prèviament o no

32. Claudicació intermitent: presència de símptomes

33. Oclusió o estenosi de >50% d'una artèria caròtida

34. Intervenció (prèvia o programada) de l'aorta abdominal, caròtides o extremitats

35. Grau funcional CCS: utilitzarem la Canadian Cardiovascular Society Classification posant el grau funcional en el moment de la indicació per a

l'angina estable i el grau funcional màxim que ha tingut el malalt en els casos en que la indicació estigui motivada per síndrome coronària aguda

36. Hipertensió pulmonar sistòlica (>60mmHg)

37. Disfunció ventricular esquerra. Es posarà la valoració per angiografia si existeix, si no, la ventriculografia, i si no l'ecocardiografia

38. Infart de miocardi. Tant si és un antecedent previ, com si és el síndrome que motiva l'ingrés, com si es produeix durant el mateix ingrés abans de la cirurgia

39. Data de l'últim infart. És important registrar la data de l'últim infart si s'ha produït durant l'últim any. Si no tenim la data exacta posar-la el més aproximada possible.

Esdeveniments durant el temps d'espera

40. Infart

41. Data de l'últim infart

42. Ingrés per angina inestable

43. Nombre d'ingressos per malaltia coronària

44. Nombre d'atencions a urgències per malaltia coronària

45. Mort

46. Data de la defunció

47. Causa cardiovascular

48. No s'opera perquè el malalt refusa la operació després d'haver estat indicada, o perquè el cirurgià refusa la operació

49. S'opera en un altre centre

50. Es produeix demora en la intervenció

51. Motiu d'aquesta demora

Motiu de l'ingrés actual

Indicarem la data i el motiu de l'ingrés en el que es realitza la intervenció. Si el malalt es trasllada d'un altre centre, posarem la data d'ingrés en el primer centre a la variable 52.

Determinants del risc en el moment de la intervenció

Tal com deiem en l'apartat de determinants del risc en el moment de la indicació, és important que la informació sobre el risc quirúrgic sigui fiable. Igualment, hem posat la opció (9) No consta, només en aquelles variables que requereixen una exploració per a confirmar-les, com és el cas de la creatinina, la oclusió de la caròtida i la hipertensió pulmonar.

58. Data de la intervenció

59. Indicació: Cirurgia urgent: el malalt s'opera durant el mateix ingrés degut a una síndrome coronària aguda; cirurgia emergent: la que es realitza durant les primeres 24 hores després de la seva indicació.

60. Última crisi d'angina de repòs en les 48h. prèvies: per tal de distingir les urgències reals d'aquelles en les que el malalt s'opera en el mateix ingrés per

criteris no clínics, s'enregistrarà si el malalt ha sofert alguna crisi d'angina de repòs durant les 48h prèvies a la intervenció.

61. Creatinina ($\leq 1.5\text{mg/dl}$)
62. Disfunció neurològica: veure variable 28
63. Malaltia hepàtica: veure variable 29
64. Malaltia pulmonar crònica: veure variable 30
65. Aneurisma ventricular esquerre
66. Claudicació intermitent
67. Oclusió o estenosi de $>50\%$ d'una artèria caròtida
68. Intervenció (prèvia o programada) de l'aorta abdominal, caròtides o extremitats

69. Grau funcional CCS: la classificació funcional està basada en la Canadian Cardiovascular Society Classification, però aplicada a malalts hospitalitzats:
 - a) posarem grau funcional IV si ha tingut angina de repòs durant les 72 hores prèvies;
 - b) grau funcional III si ha tingut angina de repòs abans de les 72 hores prèvies i es mantenen allitats o si, durant l'ingrés, han tingut angina en les activitats lleus de la hospitalització
 - c) la resta es classificaran en grau funcional II
70. Hipertensió pulmonar sistòlica ($>60\text{mmHg}$)
71. Disfunció ventricular esquerra: Veure variable 37.
72. Xoc cardiogènic
73. Angina inestable: Nitrats i.v. fins a l'entrada a quiròfan

Estat preoperatori crític:

74. Taquicardia ventricular o fibril·lació
75. Massatge cardíac preoperatori
76. Intubació abans de l'entrada a quiròfan
77. Suport ionotròpic intravenós (preoperatori)
78. Contrapulsació intraaòrtica
79. Insuf renal aguda preop.: Anúria o oligonúria $< 10\text{ml/h}$.
80. Ruptura septal postinfart

Variables relacionades amb la intervenció:

- 81-83. Nombre d'empelts de mamària, radial i/o safena, segons el cas. Posar un 0 en el que no es realitzi.
84. Cirurgia sense circulació extracorpòria

Resultats. Postoperatori immediat

85. Sagnament. Es contestarà afirmativament si es produeix un sagnament que requereixi tractament o reintervenció.
86. Infart peroperatori: quan existeixi augment de CK total o CKMB superior a 3 vegades el límit màxim normal del laboratori en les primeres 48 hores del postoperatori, associat a noves ones Q o desnivellaments nous del

segment ST \geq 1 mm. En cas d'ECG ininterpretable tenir en compte només la corba CK

87. Reintervenció

88. Complicacions infeccioses (ferida esternal, respiratòries, nosocomials). Diagnòstic clínic o bacteriològic d'infecció que justifiqui tractament antibiòtic

89. Indicar de quin tipus d'infecció es tracta

90. Altres complicacions greus cardiovasculars que hagin tingut seqüeles o han necessitat algun tractament

91. Quines

92. altres complicacions greus no cardiovasculars que hagin tingut seqüeles o han necessitat algun tractament

93. quines

94. Mort hospitalària

95. Data d'alta o defunció

Resultats. A l'any de la intervenció

96. Data de l'enquesta

97. Presència d'angina en el moment de l'enquesta

98. Grau funcional CCS: la classificació funcional està basada en la Canadian Cardiovascular Society Classification:

a) utilitzem el grau 0, no previst en la classificació original, per a tots els malalts que no tinguin angina i que duguin a terme una vida no limitada per símptomes d'altres tipus ni degut a decisió pròpia;

b) posarem el grau funcional màxim que ha tingut en el moment de l'entrevista telefònica

99. Indicar si durant el seguiment s'ha produït algun ingrés per causa cardiovascular

100. Nova revascularització amb ACTP o STENT

101. Reintervenció amb cirurgia

102. Mort

103. Data de la defunció

104. Causa cardiovascular

ANNEX 3. L'EuroSCORE i el model de l'AATRM

Escala de risc de l'AATM

Edat	70-79	7
	≥80	17
Infart recent	(<4 setmanes)	10
Grau funcional	III	4
	IV	10
Malaltia hepàtica		8
Aneurisme ventricular esquerre		11
Creatinina ≤ 1.5 mg/dl		8
Xoc cardiogènic		13
Ventilació mecànica prequirúrgica		7
Urgent o emergent		4
Reintervenció	primera	9
	Segona	15
Vàlvula mitral		6
Vàlvula tricúspide		10
Aorta toràctica		12
Valvular i coronària		7

EuroSCORE

Edat	Cada 5 anys per sobre de 60	1
Dones		1
Malatia pulmonar crònica	Ús de broncodilatadors o esteroides	1
Arteriopatia extracardíaca	Claudicació ,o oclusió o estenosi de >50% de la caròtida, o intervenció (pre o programada) de l'aorta abdominal, caòtides o extremitats	2
Disfunció neurològica	Afectació severa de l'activitat diària	2
Intervenció quirúrgica cardíaca prèvia		3
Creatinina	>200 micromol/l	2
Endocarditis activa	En tractament amb antibiòtics en el moment de la intervenció	3
Estat preoperatori crític	Taquicardia ventricular, o fibrilació, o mort sobtada abortada, o massatge cardíac preop, o ventilació preop abans de l'anestèsia, o suport inotrópic preop, o contrapulsació intraaòrtica, o insuf renal aguda preop (anúria o oligonúria < 10ml/h)	3
Angina inestable	Angina de repòs amb nitrats i.v. abans de l'anestèsia	2
Disfunció ventricular esquerra	Moderada o 30-50% Deprimida o <30	1 3
Infart recent	<90 dies	2
Hipertensió pulmonar	Sistòlica >60mmHg	2
Emergència		2
Altres intervencions a més de o en comptes de CABG		2
Intervenció aorta toràctica		3
Ruptura septal postinfart		4

ANNEX 4. Comentari editorial

González-Santos JM. ¿Gestión pública o gestión privada? Implicaciones en los resultados de la cirugía coronaria. Rev Esp Cardiol 2006;59:414-7

González Santos JM

¿Gestión pública o gestión privada? Implicaciones en los resultados de la cirugía coronaria

Revista Española de Cardiología 2006; 59(5):414-7

