Open Access: free full text available online
For the full versions of these articles see bmj.com

RESEARCH

Impact of pulse oximetry screening on the detection of duct dependent congenital heart disease: a Swedish prospective screening study in 39 821 newborns

Anne de-Wahl Granelli,[1,11] Margareta Wennergren,[2] Kenneth Sandberg,[3] Mats Mellander,[1] Carina Bejlum,[4] Leif Inganäs,[5] Monica Eriksson,[6] Niklas Segerdahl,[7] Annelie Ågren,[8] Britt-Marie Ekman-Joelsson,[9] Jan Sunnegårdh,[1] Mario Verdicchio,[10] Ingegerd Östman-Smith[1,11]

EDITORIAL by Barrington

[1]Department of Paediatric Cardiology, Queen Silvia Children's Hospital, S-416 85 Göteborg, Sweden
[2]Department of Obstetrics, Sahlgrenska University Hospital, S-416 85 Göteborg
[3]Department of Neonatology, Queen Silvia Children's Hospital
[4]Department of Obstetrics, NÄL Hospital, S-461 85 Trollhättan, Sweden
[5]Children's Department, NÄL Hospital
[6]Obstetric Department, Södra Älvsborgs Hospital, S-501 15 Borås, Sweden
[7]Children's Department, Södra Älvsborgs Hospital
[8]Department of Neonatology, Hospital of Skövde, S-541 85 Skövde, Sweden
[9]Children's Department, Hospital of Skövde
[10]Institute of Forensic Medicine, S-405 30 Göteborg
[11]Institute of Clinical Sciences, Sahlgrenska Academy, Gothenburg University, S-405 30 Göteborg

Correspondence to:
I Östman-Smith
ingegerd.ostman-smith@pediat.gu.se

Cite this as: *BMJ* 2009;338:a3037
doi:10.1136/bmj.a3037

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;338:a3037

ABSTRACT

Objective To evaluate the use of pulse oximetry to screen for early detection of life threatening congenital heart disease.

Design Prospective screening study with a new generation pulse oximeter before discharge from well baby nurseries in West Götaland. Cohort study comparing the detection rate of duct dependent circulation in West Götaland with that in other regions not using pulse oximetry screening. Deaths at home with undetected duct dependent circulation were included.

Setting All 5 maternity units in West Götaland and the supraregional referral centre for neonatal cardiac surgery.

Participants 39 821 screened babies born between 1 July 2004 and 31 March 2007. Total duct dependent circulation cohorts: West Götaland n=60, other referring regions n=100.

Main outcome measures Sensitivity, specificity, positive and negative predictive values, and likelihood ratio for pulse oximetry screening and for neonatal physical examination alone.

Results In West Götaland 29 babies in well baby nurseries had duct dependent circulation undetected before neonatal discharge examination. In 13 cases, pulse oximetry showed oxygen saturations ≤90%, and (in accordance with protocol) clinical staff were immediately told of the results. Of the remaining 16 cases, physical examination alone detected 10 (63%). Combining physical examination with pulse oximetry screening had a sensitivity of 24/29 (82.8% (95% CI 64.2% to 95.2%)) and detected 100% of the babies with duct dependent lung circulation. Five cases were missed (all with aortic arch obstruction). False positive rate with pulse oximetry was substantially lower than that with physical examination alone (69/39 821 (0.17%) *v* 729/38 413 (1.90%), P<0.0001), and 31/69 of the "false positive" cases with pulse oximetry had other pathology. Thus, referral of all cases with positive oximetry results for echocardiography resulted in only 2.3 echocardiograms with normal cardiac findings for every true positive case of duct dependent circulation. In the cohort study, the risk of leaving hospital with undiagnosed duct dependent circulation was 28/100 (28%) in other referring regions *v* 5/60 (8%) in West Götaland (P=0.0025, relative risk 3.36 (95% CI 1.37 to 8.24)). In the other referring regions 11/25 (44%) of babies with transposition of the great arteries left hospital undiagnosed versus 0/18 in West Götaland (P=0.0010), and severe acidosis at diagnosis was more common (33/100 (33%) *v* 7/60 (12%), P=0.0025, relative risk 2.8 (1.3 to 6.0)). Excluding premature babies and Norwood surgery, babies discharged without diagnosis had higher mortality than those diagnosed in hospital (4/27 (18%) *v* 1/110 (0.9%), P=0.0054). No baby died from undiagnosed duct dependent circulation in West Götaland versus five babies from the other referring regions.

Conclusion Introducing pulse oximetry screening before discharge improved total detection rate of duct dependent circulation to 92%. Such screening seems cost neutral in the short term, but the probable prevention of neurological morbidity and reduced need for preoperative neonatal intensive care suggest that such screening will be cost effective long term.

INTRODUCTION

About 1-1.8 babies per 1000 live births have a duct dependent circulation, with a persistent ductus arteriosus being necessary for survival.[1-3] The effects of ductal closure may not be apparent at early discharge examination from maternity units. In the UK 30% of babies with critical heart disease leave hospital undiagnosed,[1] and in Sweden the proportions of missed cases have increased over the past decade (see bmj.com for discussion of the possible reasons for this).[4]

Screening infants with non-invasive measurement of oxygen saturation has been proposed as an aid for early detection of duct dependent circulation. We evaluated strategies to maximise sensitivity while minimising false positives with a new generation oximeter which measured functional oxygen saturation preductally (in right hand) and postductally (in either foot). We arrived at optimal screening cut-off values of <95% saturation or >3% difference between right hand and foot.[5] We also found that the type of oximeter used had a significant effect on both the detection rate and false positive rate.[5]

Using these cut-off values and a new generation oximeter, we have now conducted a large, prospective, multicentre study of routine screening with pulse oximetry in the well baby units in the West Götaland region of Sweden. Our new study incorporates further strategies to reduce false positive results from pulse oximetry screening. See bmj.com.

METHODS

Study population

Prospective pulse oximetry screening of all babies in well baby nurseries in West Götaland began with a sequential rolling start in all the five maternity units between July and November 2004 and lasted to the end of March 2007.

Cohort population

We included all babies born with duct dependent circulation in West Götaland (total live births=46 963) and in the other referring regions (total live births=108 604). We excluded from comparison those babies with a prenatal diagnosis of duct dependent circulation.

Screening study

Prospective screening of oxygen saturation was conducted preductally (palm of right hand) and postductally (either foot) with identical pulse oximeters on all newborn infants before routine neonatal physical examination.

In order for the screening of all babies to be logistically feasible and to take no more than five minutes of nursing time, it had to be incorporated in ordinary nursing routines and was usually carried out by trained midwives, nurses, and nursery nurses before the daily weighing that preceded the discharge examination.

Study protocol

When both preductal and postductal oxygen saturation was <95% or the difference between the two measurements was ≥3% (≥2 standard deviations of inter-observer measurement variability[5]) the baby was provisionally considered to be screening positive, but a repeat measurement was performed. Babies with three repeated positive measurements were supposed to have an echocardiogram performed the same day according to the study protocol, but some babies scheduled for early discharge had only two pulse oximetry screenings.

Neonatal physical examination—After the routine physical examination, the examining paediatrician recorded (*a*) no suspicion of congenital heart defect, (*b*) weak suspicion of congenital heart disease, or (*c*) strong evidence of congenital heart defect, and had to state whether referral for echocardiography was indicated based on the physical findings. This was recorded before the paediatrician was shown the results of the pulse oximetry screening.

Cohort study

We compared the overall rate of detection of duct dependent circulation in West Götaland with that in other regions not using pulse oximetry screening but which also refer children to the same supra-regional centre for congenital cardiac surgery. We examined the medical records of all babies with duct dependent circulation in the two cohort populations and recorded preoperative acidosis and 30 day mortality.

We retrieved data from the national database of the National Board of Forensic Medicine, for information of all deaths due to undiagnosed cardiovascular malformations in children under 1 year of age in Sweden born during the study period. We compared the number of deaths from undiagnosed duct dependent circulation (which all occurred within 30 days of birth) in West Götaland with that in the other referring regions.

Statistical analysis

We calculated sensitivity, specificity, positive predictive value, and negative predictive values for pulse oximetry screening and for blind neonatal physical examination alone.

RESULTS

Screening study

Of the 39 899 newborns eligible for the screening study, 39 821 (99.8%) had completed the pulse oximetry protocols and 38 429 (96.3%) had complete data from both pulse oximetry and physical examination (see bmj.com).

Pulse oximetry screening

Details of the 29 babies in the screening study found to have duct dependent circulation, including the results from pulse oximetry screening and the physical examination are shown on bmj.com. Pulse oximetry in isolation gave abnormal screening results in 19/29 (66%) of apparently well babies with duct dependent circulation.

The sensitivity of the pulse oximetry for detecting pulmonary duct dependent circulation and transposition of the great arteries was 9/9, but the sensitivity for essentially acyanotic left heart obstruction was, unsurprisingly, lower (10/20). However, one child with positive pulse oximetry result and interrupted aortic arch was discharged home without echocardiography in violation of the study protocol, and thus the real life sensitivity of the pulse oximetry screening was 18/29 (62%). The table shows the sensitivity, specificity, positive and negative predictive values, and likelihood ratio for pulse oximetry. A positive pulse oximetry screening gives a relative risk of 719.8 (95% confidence interval 350.3 to 1479; P<0.0001) of having duct dependent heart disease.

In accordance with the study protocol, the examining neonatologist was immediately informed of the pulse oximetry results for the 12 babies with oxygen saturation ≤90%, and for one with saturation of 91%. Thus, these 13 were excluded from the evaluation of neonatal physical examination alone.

In terms of which screening criteria were positive, both preductal and postductal oxygen saturations were <95% in 13 babies, while in five babies a difference of >3% between preductal and postductal saturations was the only positive criterion. However, many babies with

complex cyanotic heart disease were positive on both criteria, so in total 14 babies had a saturation difference >3%, of whom eight (57%) had duct dependent systemic circulation. Babies without critical congenital heart defect or lung pathology had a median oxygen saturation of 99% (interquartile range 98% to 100%) both preductally and postductally. The median age at screening was 38 hours (interquartile range 5.5 to 95.5, range 1 to 406), and 90% of the babies were screened at ≤72 hours of age. The earliest permitted discharge in West Götaland is 6 hours after birth, and 1317 babies (3.3%) were screened that early.

Neonatal physical examination alone

Physical examination alone detected 10/16 cases of duct dependent circulation (that is, a sensitivity of 63%) (table). The positive predictive value of neonatal examination was significantly lower than that of pulse oximetry and the likelihood ratio was lower.

The clinical findings from physical examination that provoked referral for echocardiography in the 10 detected babies with duct dependent circulation were systolic murmurs (n=5), poor or absent femoral pulses with a murmur (n=4), and poor femoral pulses alone (n=1). Thus poor or absent femoral pulses was the alerting sign in half of these children, and contributed to the detection of two babies with duct dependent systemic circulation who would otherwise have been missed with oximetry screening.

Combining pulse oximetry and physical examination

As different cases were missed by clinical examination and by pulse oximetry, the combination of neonatal physical examination and oximetry screening had a higher sensitivity than either of the methods individually, although the higher number of false positives from physical examination lowered the positive predictive value to 2.92 (1.98 to 4.31) (see table). Of the infants detected by oximetry, at least four had no physical signs that would have led to a referral for echocardiography. Five babies with duct dependent circulation were discharged without diagnosis during the study; all had duct dependent systemic circulation due to aortic arch obstruction, but, as one was discharged with a positive oximetry result in violation of study protocol, the potential detection rate in apparently well babies was 86% (25/29).

False positive results with pulse oximetry

The "false" positive rate for oximetry screening was 69/39 821 (0.17%). Of these, 45% (31/69) had other significant heart malformation, lung problem, or infection. Only 41 babies with positive oximetry results had normal cardiac findings on echocardiography. Thus there were 2.3 echocardiograms with normal findings per baby with duct dependent heart disease detected by pulse oximetry screening (41/18). See bmj.com for details.

False positive results with neonatal physical examination

Physical examination alone generated 739 referrals for echocardiography with a false positive rate (729/38 374 (1.91%)) more than 10 times higher than that for pulse oximetry (P<0.0001). See bmj.com.

Cohort population

Between 1 July 2004 and 31 March 2007, the birth prevalence of duct dependent circulation in West Götaland was 62/46 963 (1.32/1000), two with prenatal diagnosis. In all other referring regions, not using pulse oximetry screening but some with prenatal screening by echocardiography, 109/108 604 newborn infants had duct dependent circulation (birth prevalence 1.00/1000). Of these, 100 were included in our comparison, as nine had prenatal diagnosis.

The risk of leaving hospital with undiagnosed duct dependent circulation was 28/100 (28%) in the other referring regions versus 5/60 (8%) in West Götaland (P=0.0025; relative risk 3.36 (95% confidence interval 1.37 to 8.24)). The difference was mainly because of the improved detection of pulmonary duct dependent circulation (where we included transposition of the

The performance of screening methods in the detection of duct dependent circulation in newborn infants in West Götaland (1 July 2004 to 31 March 2007)

Performance	Physical examination alone (n=38 374)	Pulse oximetry (n=38 429)	Physical examination plus pulse oximetry (n=38 429)
Sensitivity (95% CI) (%)	62.50 (35.43 to 84.80)*	62.07 (42.3 to 79.31)	82.76 (64.23 to 94.15)
Specificity (95% CI) (%)	98.07 (97.93 to 98.21)	99.82 (99.77 to 99.86)	97.88 (97.73 to 98.03)
Positive predictive value (95% CI) (%)	1.35 (0.65 to 2.47)	20.69 (12.75 to 30.71)	2.92 (1.88 to 4.31)
Negative predictive value (95% CI) (%)	99.98 (99.96 to 99.99)	99.97 (99.95 to 99.99)	99.99 (99.97 to 100.00)
Likelihood ratio	32.37	344.8	39.08
False positive rate (%)	1.90	0.17†	2.09
No of true positives	10*	18‡	24‡
No of false negatives	6*	11§	5§
No of false positives	729	69	798
No of true negatives	37 022	38 259	36 881
Relative risk (95% CI) (%)	83.6 (30.5 to 229.5)	719.8 (350.3 to 1479)	215.4 (82.4 to 563.0)

*Blind physical examination alone cannot be compared directly with the other two methods as the number of babies with duct dependent circulation was 16 in this group.
†False positive rate calculated on total numbers of patients completing pulse oximetry (n=39 821).
‡Patient who was diagnosed after repeated failures of obtaining a pulse oximetry signal in the feet is counted as true positive.
§Patient who fulfilled screening criteria but was discharged due to protocol violation is counted as false negative.

WHAT IS ALREADY KNOWN ON THIS TOPIC

About 1-2 babies per 1000 live births have an immediately life threatening cardiac malformation, and 30% of such infants leave hospital without the malformation being recognised and either return to hospital in circulatory collapse or die at home

Pulse oximetry screening has been advocated as a possible tool to improve detection, but sensitivity and cost effectiveness remain unproved in the absence of sizeable prospective studies

WHAT THIS STUDY ADDS

As inpatient maternity stays have reduced, an increasing proportion of babies with duct dependent pulmonary circulation leave hospital undetected

Pulse oximetry screening performed both preductally and postductally detects 100% of infants with pulmonary duct dependent circulation and, when combined with routine clinical examination, detects 92% of all infants with duct dependent circulation before hospital discharge, and has a higher detection rate than physical examination alone

Introduction of pulse oximetry screening is cost neutral in the immediate perspective, as each additional case that receives a timely diagnosis costs the same as the treatment of a child that

great arteries) in West Götaland (odds ratio 18.83 (1.07 to 331), P=0.0030). Among the 12 babies with duct dependent pulmonary circulation who had been discharged home (all in the regions other than West Götaland), 11 had transposition of the great arteries. Severe acidosis at diagnosis was more common among unscreened babies (33/100 (33%) *v* 7/60 (12%), P=0.0025, relative risk 2.8 (1.3 to 6.0)). Excluding premature babies and Norwood surgery, babies discharged without diagnosis had higher mortality than those diagnosed in hospital (4/27 (18%) *v* 1/110 (0.9%), P=0.0054).

No children with undiagnosed duct dependent circulation died in West Götaland (0/60) compared with 5/100 in the other referring regions (P=0.16). See bmj.com for further details.

DISCUSSION

Principal findings

In asymptomatic babies we found that the combination of neonatal physical examination plus pulse oximetry screening for duct dependent heart disease had a detection rate of 82.8% (86.2% if protocol violations are ignored), with a low false positive rate of 0.17% for pulse oximetry. Because of the large sample size our estimate provides an authoritative assessment of this screening method. However, about half of the babies with duct dependent disease presented clinically before discharge examination, so that in total the introduction of pulse oximetry screening meant that in our region of West Götaland 92% of all babies with a duct dependent circulation were diagnosed before leaving hospital. This is a significantly higher proportion than that encountered among babies from other Swedish regions not using pulse oximetry screening (72%; P=0.0025).

The detection rate of blind physical examination alone was 62.5%. In the region using pulse oximetry screening there were no deaths in the community from undiagnosed critical heart disease, but there were five deaths, 5% of babies with duct dependent circulation, in the regions not using pulse oximetry screening. This improved detection was achieved by an alteration of nursing routines, that was estimated to increase nursing time spent per baby by maximum five minutes, and occasioned only 2.3 extra echocardiograms without pathology per case of true positive duct dependent heart disease detected by pulse oximetry.

There is no routine fetal echocardiography in our region, leading to a low rate of antenatal detection of duct dependent heart disease (3.3%), much lower than the nearly 20% antenatal detection of all critical heart disease over the last few years in Newcastle.[1]

Strengths and weaknesses of our study

The major strengths of our study are the large number of babies prospectively screened, and use of the Swedish personal identity number system together with the forensic database, so that we can be certain that no deaths in the community or elsewhere have been overlooked in the screened cohort.

A weakness of our study design was that it was impossible for ethical reasons to withhold seriously deranged pulse oximetry values from the attending medical staff, which meant that our evaluation of the success of physical examination alone to detect duct dependent heart disease excluded the most severely cyanotic types of duct dependent disease. This is the main reason why we included a contemporary comparison cohort from the other Swedish regions. This comparison group showed that, without pulse oximetry screening, 23% of patients even with cyanotic duct dependent pulmonary circulation left hospital undiagnosed, and there was no significant difference in the other referring regions between the number of cases with pulmonary versus systemic duct dependent circulation that was missed (P=0.62). See bmj.com.

Results in relation to other studies

Previously published studies attempting to assess the potential of pulse oximetry for the screening for critical congenital heart disease have been too small to enable a confident estimate of sensitivity because of the prevalence of such disease being only 1-1.8/1000.[2 3 6-10] Depending on the cut-off criteria, the false positive rate of pulse oximetry screening varied between 0.009% and 5% in these studies. Richmond et al showed that the introduction of repeat pulse oximetry brought their false positive rate down from 5% to 1%.[2]

The detection rate of physical examination alone, 62.5% in our study, agrees with the 62% postnatal in-hospital detection rate reported by Wren et al with on average 30% of their babies with critical heart disease leaving hospital undiagnosed.[1]

Few of the earlier studies of pulse oximetry screening have compared it with the detection rate of physical examination alone in well babies. See bmj.com.

A large Norwegian multicentre study published after we submitted our manuscript used the same pulse oximeter and probes as in our study, but the participants measured only postductal oxygen saturation with a cut-off point of <95% in two repeated measurements, which would have had a sensitivity of

60.7% in our population(see bmj.com).[11] We maintain that our protocol is preferable.

Cost benefit analysis

Using Griebsch et al's detailed model and their highest cost estimate for echocardiography,[12] we calculate from our screening results that the cost for 18 timely diagnoses made by pulse oximetry is £3430 (€3785; $5140) per timely diagnosis made. As the cost for an infant leaving hospital with duct dependent circulation and returning in circulatory collapse was calculated to be £3453,[12] the introduction of pulse oximetry screening should be, at a minimum, cost neutral since each additional case diagnosed saves at least as much as each missed case costs.

The cost per timely diagnosis for clinical examination in our study came out between £7700 (for those actually referred for echocardiography from physical examination alone) and £2526 (in the unlikely event that all the infants with pulse oximetry results of ≤90% saturation would have been referred to echocardiography from physical examination alone).

As well as the acute costs, a timely diagnosis improves the survival of affected babies and reduces possible long term neurological morbidity secondary to circulatory collapse (see bmj.com).

We thank everyone, especially nursery and midwifery staff, who contributed to this multicentre study; Peter Skoog, of the department of biomedical engineering, for invaluable help with the pulse oximeters; and Kajsa Boll, Britt-Marie Carlsson, Anna-Maria Edwardsson, Anna Larsson, Josefin Mellander, and Niklas Smith for help with data entry.
Contributors: See bmj.com.
Ewa-Lena Bratt conducted the pilot period at Östra together with AWG.
Funding: Grants from the Regional Science and Education Foundation, West Götaland Region, supporting inter-regional collaboration, and a LUA-ALF project grant from Gothenburg University.
Competing interests: None declared.
Ethical approval: The Gothenburg University Research and Ethics Committee approved this study (application No Ö670-03).
Provenance and peer review: Not commissioned.

1 Wren C, Reinhardt Z, Khawaja K. Twenty-year trends in diagnosis of life-threatening neonatal cardiovascular malformations. *Arch Dis Child Fetal Neonatal Ed* 2008;93:F33-5.
2 Richmond S, Reay G, Abu Harb M. Routine pulse oximetry in the asymptomatic newborn. *Arch Dis Child Fetal Neonatal Ed* 2002;87:F83-8.
3 Koppel RI, Druschel CM, Carter T, Goldberg BE, Mehta PN, Talwar R, et al. Effectiveness of pulse oximetry screening for congenital heart disease in asymptomatic newborns. *Pediatrics* 2003;111:451-5.
4 Mellander M, Sunnegardh J. Failure to diagnose critical heart malformations in newborns before discharge—an increasing problem? *Acta Paediatr* 2006;95:407-13.
5 De-Wahl Granelli A, Mellander M, Sunnegardh J, Sandberg K, Östman-Smith I. Screening for duct-dependent congenital heart disease with pulse oximetry: a critical evaluation of strategies to maximize sensitivity. *Acta Paediatr* 2005;94:1590-6.
6 Hoke TR, Donohue PK, Bawa PK, Mitchell RD, Pathak A, Rowe PC, et al. Oxygen saturation as a screening test for critical congenital heart disease: a preliminary study. *Pediatr Cardiol* 2002;23:403-9.
7 Reich JD, Miller S, Brogdon B, Casatelli J, Gompf TC, Huhta JC, et al. The use of pulse oximetry to detect congenital heart disease. *J Pediatr* 2003;142:268-72.
8 Bakr AF, Habib HS. Combining pulse oximetry and clinical examination in screening for congenital heart disease. *Pediatr Cardiol* 2005;26:832-5.
9 Rosati E, Chitano G, Dipaola L, De Felice C, Latini G. Indications and limitations for a neonatal pulse oximetry screening of critical congenital heart disease. *J Perinat Med* 2005;33:455-7.
10 Arlettaz R, Bauschatz AS, Monkhoff M, Essers B, Bauersfeld U. The contribution of pulse oximetry to the early detection of congenital heart disease in newborns. *Eur J Pediatr* 2006;165:94-8.
11 Meberg A, Brugmann-Pieper S, Due R Jr, Eskedal L, Fagerli I, Farstad T, et al. First day of life pulse oximetry screening to detect congenital heart defects. *J Pediatr* 2008;152:761-5.
12 Griebsch I, Knowles RL, Brown J, Bull C, Wren C, Dezateux CA. Comparing the clinical and economic effects of clinical examination, pulse oximetry, and echocardiography in newborn screening for congenital heart defects: a probabilistic cost-effectiveness model and value of information analysis. *Int J Technol Assess Health Care* 2007;23:192-204.

Accepted: 5 October 2008

Effect of fish oil on arrhythmias and mortality: systematic review

Hernando León,[1,2] Marcelo C Shibata,[1,2] Soori Sivakumaran,[1,2] Marlene Dorgan,[3] Trish Chatterley,[3] Ross T Tsuyuki[1,2]

EDITORIAL by Brunner and Iso

[1]Epidemiology Coordinating and Research (EPICORE) Centre, Edmonton, AB, Canada
[2]Department of Medicine, Division of Cardiology, Faculty of Medicine and Dentistry, University of Alberta, 220 College Plaza, Edmonton, AB, Canada T6G 2C8
[3]Institute of Health Economics and John W Scott Health Sciences Library, University of Alberta
Correspondence to: R T Tsuyuki
ross.tsuyuki@ualberta.ca

Cite this as: *BMJ* 2008;337:a2931
doi:10.1136/bmj.a2931

ABSTRACT

Objective To synthesise the literature on the effects of fish oil—docosahexaenoic acid (DHA) and eicosapentaenoic acid (EPA)—on mortality and arrhythmias and to explore dose response and formulation effects.

Design Systematic review and meta-analysis.

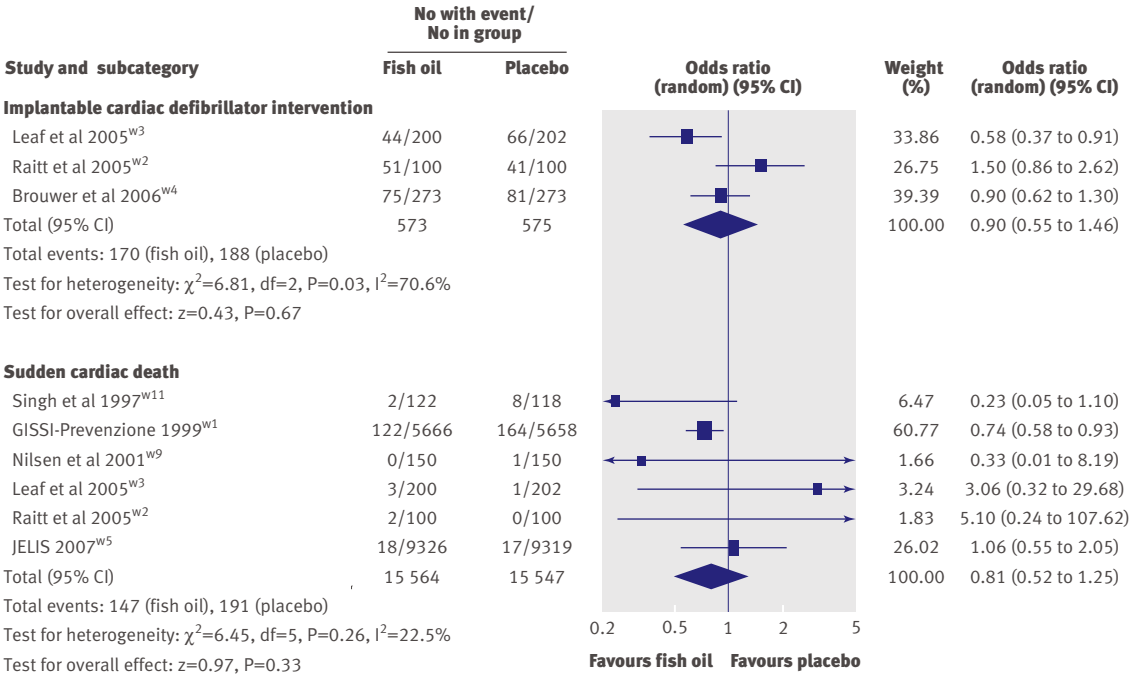
Data sources Medline, Embase, the Cochrane Library, PubMed, CINAHL, IPA, Web of Science, Scopus, Pascal, Allied and Complementary Medicine, Academic OneFile, ProQuest Dissertations and Theses, Evidence-Based Complementary Medicine, and LILACS.

Studies reviewed Randomised controlled trials of fish oil as dietary supplements in humans.

Data extraction The primary outcomes of interest were the arrhythmic end points of appropriate implantable cardiac defibrillator intervention and sudden cardiac death. The secondary outcomes were all cause mortality and death from cardiac causes. Subgroup analyses included the effect of formulations of EPA and DHA on death from cardiac causes and effects of fish oil in patients with coronary artery disease or myocardial infarction.

Data synthesis 12 studies totalling 32 779 patients met the inclusion criteria. A neutral effect was reported in three studies (n=1148) for appropriate implantable cardiac defibrillator intervention (odds ratio 0.90, 95% confidence interval 0.55 to 1.46) and in six studies (n=31 111) for sudden cardiac death (0.81, 0.52 to 1.25). 11 studies (n=32 439 and n=32 519) provided data on the effects of fish oil on all cause mortality (0.92, 0.82 to 1.03) and a reduction in deaths from cardiac causes (0.80, 0.69 to 0.92). The dose-response relation for DHA and EPA on reduction in deaths from cardiac causes was not significant.

Conclusions Fish oil supplementation is associated with a significant reduction in deaths from cardiac causes but



Effect of fish oil on implantable cardiac defibrillator intervention and sudden cardiac death

had no effect on arrhythmias or all cause mortality. Evidence to recommend an optimal formulation of EPA or DHA to reduce these outcomes is insufficient. Fish oils are a heterogeneous product, and the formulations for DHA and EPA remain unclear.

INTRODUCTION

An interest in omega 3 fats (fish oil) to prevent and treat cardiovascular diseases arose from a report of a decreased risk of cardiovascular disease in Inuit with a high intake of fish oil.[1] The GISSI-Prevenzione trial[w1] of 11 324 patients randomised to a mixture of the omega 3 fats eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) or placebo showed a significant reduction in all cause mortality and death from cardiovascular causes over 3.5 years of follow-up. These results seemed to be driven by a reduction in sudden cardiac death, sparking an interest in the potential antiarrhythmic properties of fish oil. Three recent randomised clinical trials looked at fish oil in the prevention of sudden cardiac death in patients with implantable cardiac defibrillators.[w2-w4] None of these trials, or a recent systematic review,[2] showed a beneficial effect of fish oil on patient outcomes. This may be because of methodological limitations such as different formulations of fish oil. We evaluated the effect of EPA and DHA on all cause mortality and deaths from cardiac causes based on the formulation of these compounds and systematically examined the association between fish oil and arrhythmic events.

METHODS

We searched 15 databases (see bmj.com) for randomised controlled trials of fish oil as dietary supplements in humans. The primary outcomes of interest were the arrhythmic end points of implantable cardiac defibrillator intervention and sudden cardiac death. The secondary outcomes were all cause mortality and death from cardiac causes. Subgroup analyses included the effect of formulations of EPA and DHA on deaths from cardiac causes and effects of fish oil in patients with coronary artery disease or myocardial infarction.

Two investigators independently reviewed selected abstracts, with disagreements resolved by a third reviewer. The quality of each study was assessed by the Jadad criteria.[3]

Statistical analysis

Data were entered into 2×2 tables and analysed using RevMan 4.2.9. Data are presented as odds ratios with 95% confidence intervals, using the DerSimonian and Laird random effects model. For evaluation of heterogeneity we used the χ^2 and I^2 tests. The z test was used to determine overall effect. Assessment of publication bias was carried out by generating a funnel plot from the end points showing a significant benefit from fish oil.

To evaluate whether a dose-response exists with EPA or DHA we carried out a meta-regression analysis using the random effects model for meta-analysis regression, and analysed the data using STATA. For this analysis we used the outcome of deaths from cardiac causes.

We evaluated the relative risk of a non-cardiovascular adverse effect occurring in patients who received fish oil compared with placebo. We also expressed treatment effects and adverse effects (see bmj.com) using the numbers needed to treat to reduce one significant clinical event and numbers needed to harm for an adverse effect.

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a2931

RESULTS

After exclusions, 12 studies were included (see bmj.com).[w1-w12] Five scored 5 for methodological quality on the Jadad scale, four scored 4, two scored 3, and one scored 2.

In three studies (n=1148) fish oil supplementation reduced the risk of implantable cardiac defibrillator intervention by a non-significant 10% (odds ratio 0.90, 95% confidence interval 0.55 to1.46; figure). Six studies (n=31 111) evaluated the effect of fish oil on the incidence of sudden cardiac death (figure). A non-significant reduction was observed (odds ratio 0.81, 0.52 to 1.25).

Eleven studies (n=32 519) showed a significant 20% decrease in death from cardiac causes (odds ratio 0.80, 0.69 to 0.92; see bmj.com). The funnel plot, however, suggests some publication bias (see bmj.com). Eleven studies evaluated all cause mortality (n=32 439) and showed a non-significant 8% reduction (odds ratio 0.92, 0.82 to 1.03). These results were driven primarily by the GISSI-Prevenzione and JELIS trials (see bmj.com).[w1 w5]

A meta-regression analysis did not show a dose-response relation between DHA and EPA and death from cardiac causes.

A subgroup analysis was carried out on the effect of fish oil on sudden cardiac death and death from cardiac causes in patients with coronary artery disease or after myocardial infarction, including the secondary prevention arm of the JELIS trial.[w5] Four studies (n=15 528) showed a 26% reduction (0.74, 0.59 to 0.92) in sudden cardiac death with fish oil (see bmj.com). Eight studies (n=16 390) showed a significant 20% reduction in deaths from cardiac causes compared with placebo (0.80, 0.69 to 0.93; see bmj.com).

The incidence of adverse effects was 10.5% in patients receiving fish oil compared with 6.7% in those receiving placebo. Patients taking fish oil are therefore 56% more likely to experience an adverse effect than those taking placebo; most of these effects were described as mild. This corresponds to a number needed to treat with fish oil to reduce one death from cardiac causes of 189, and a number needed to harm of 26.

DISCUSSION

Our systematic review of randomised trials of fish oil supplementation showed no beneficial effect on arrhythmic events or all cause mortality but a significant reduction in deaths from cardiac causes. This is in contrast to the results of the GISSI-Prevenzione trial,[w1] which suggested that the beneficial effect on death was due primarily to a reduction in sudden cardiac death.

This is the first systematic review attempting to evaluate whether the protective mechanism of fish oil supplementation is related to a reduction of arrhythmic episodes determined by a reduction in either implantable cardiac defibrillator interventions or sudden cardiac death. We found a neutral effect on these outcomes. The confidence intervals for these outcomes were wide and a beneficial effect up to a 45-48% relative risk reduction cannot be excluded. Some heterogeneity was found among the three studies that assessed implantable cardiac defibrillator intervention (see figure). Also, regardless of the high quality of these studies, differences among them were substantial, including dosages of fish oil and study power.

Our analysis showed a highly variable effect of fish oil supplementation on sudden cardiac death, varying from a 48% reduction in events to a 25% increase (figure). These results were mainly driven by two studies, the GISSI-Prevenzione and JELIS trials.[w1 w5] GISSI-Prevenzione was a secondary prevention study in patients after myocardial infarction with a moderate fish intake and mainly DHA supplementation,[w1] whereas the JELIS trial included patients with hypercholesterolaemia for both primary and secondary prevention who had a high fish intake and received statins and only EPA supplementation.[w5]

We observed a significant 20% reduction in deaths from cardiac causes. This is similar to systematic reviews that included studies of dietary intake of fish oil by increased intake or supplementation, which showed a significant reduction in cardiovascular events.

We did not find a significant reduction in all cause mortality, unlike previous reviews.[4-7] The upper limit of the confidence interval was only 1.03, which might represent a lack of power for this end point. We also excluded studies on dietary supplementation.

We did not observe a dose-response relation for effect on deaths from cardiac causes according to dosage of EPA or DHA.

Limitations

The results of this systematic review were driven primarily by two large, but different trials,[w1 w5] which accounted for 92% of the patients. Nevertheless, our review was the first to include the JELIS trial, totalling more than 18 000 patients.[w5] Despite our comprehensive search strategy, the funnel plot suggested publication bias.

We also found a wide variability (0-2000 mg/day) in amounts of EPA and DHA in the formulations among the studies, making it difficult to determine the optimal dosage.

Another limitation in our analysis was heterogeneity among the outcomes measured. Significant heterogeneity (I^2=70.6%) was found among the implantable cardiac defibrillator trials, which might affect the validity of this analysis (see figure). The heterogeneity for the other outcomes measured in our study are, however, relatively low.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Fish oils are thought to be associated with a reduction in deaths from cardiac causes

Systematic reviews have been inconclusive and did not include recent studies

WHAT THIS STUDY ADDS

Fish oils had no significant effect on reduction in arrhythmic events but were associated with a significant reduction in deaths from cardiac causes

No evidence was found of a dose-response relation between type of fish oil and reduction in deaths from cardiac causes

Our analysis of appropriate implantable cardiac defibrillator intervention included 1148 patients from secondary prevention trials and showed a non-significant effect of fish oil that might be a type 2 error. Therefore our pooled analysis for this end point should be considered inconclusive on the basis of small sample size and noticeable heterogeneity among the trials.

Implications for clinical practice

The results of this systematic review show a beneficial effect of fish oil in reducing deaths from cardiac causes. The optimal dose or formulation of fish oil is unknown, but it seems reasonable to use a daily formulation similar to that in the GISSI-Prevenzione trial (465 mg EPA/386 mg DHA). The effect of fish oil on arrhythmic events remains inconclusive.

We thank Scot Simpson for his help with Rev Man Software, Jeppe Hagstrup Christensen (Aalborg Hospital, Aarhus University Hospital, Denmark) for sharing his results, and Ben Vandermeer for his help with the meta-regression analysis. No compensation of any kind was received by them for their help in preparing this manuscript. HL is a recipient of a research fellowship from the Alberta Heritage Foundation for Medical Research. RTT is supported by the University of Alberta Merck Frosst chair in patient health management.

Contributors: See bmj.com.
Funding: None.
Competing interests: None declared.
Ethical approval: Not required.
Provenance and peer review: Not commissioned; externally peer reviewed.

1 Bang HO, Dyerberg J, Hjoorne N. The composition of food consumed by Greenland eskimos. *Acta Med Scand* 1976;200:69-73.
2 Jenkins DJ, Josse AR, Beyene J, Dorian P, Burr ML, LaBelle R, et al. Fish-oil supplementation in patients with implantable cardioverter defibrillators: a meta-analysis. *CMAJ* 2008;178:157-64.
3 Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;17:1-12.
4 Hooper L, Thompson RL, Harrison RA, Summerbell CD, Ness AR, Moore HJ, et al. Risks and benefits of omega 3 fats for mortality, cardiovascular disease, and cancer: systematic review. *BMJ* 2006;332:752-60.
5 Yzebe D, Lievre M. Fish oils in the care of coronary heart disease patients: a meta-analysis of randomized controlled trials. *Fundam Clin Pharmacol* 2004;18:581-92.
6 Mozaffarian D, Rimm EB. Fish intake, contaminants, and human health: evaluating the risks and the benefits. *JAMA* 2006;296:1885-99.
7 Bucher HC, Hengstler P, Schindler C, Meier G. N-3 polyunsaturated fatty acids in coronary heart disease: a meta-analysis of randomized controlled trials. *Am J Med* 2002;112:298-304.

Accepted: 1 October 2008

Reporting and interpretation of SF-36 outcomes in randomised trials: systematic review

Despina G Contopoulos-Ioannidis,[1,2] Anastasia Karvouni,[3] Ioanna Kouri,[3] John P A Ioannidis[3,4]

EDITORIAL by Garratt

[1]Department of Paediatrics, University of Ioannina School of Medicine, Ioannina, Greece
[2]Department of Paediatrics, George Washington University, School of Medicine and Health Sciences, Washington, DC, USA
[3]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Greece
[4]Institute for Clinical Research and Health Policy Studies, Tufts University School of Medicine, Boston, MA, USA
Correspondence to: J P A Ioannidis, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina 45110, Greece
jioannid@cc.uoi.gr

Cite this as: *BMJ* 2009;338:a3006
doi:10.1136/bmj.a3006

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009:338:a3006

ABSTRACT

Objective To determine how often health surveys and quality of life evaluations reach different conclusions from those of primary efficacy outcomes and whether discordant results make a difference in the interpretation of trial findings.
Design Systematic review.
Data sources PubMed, contact with authors for missing information, and author survey for unpublished SF-36 data.
Study selection Randomised trials with SF-36 outcomes (the most extensively validated and used health survey instrument for appraising quality of life) that were published in 2005 in 22 journals with a high impact factor.
Data extraction Analyses on the two composite and eight subdomain SF-36 scores that corresponded to the time and mode of analysis of the primary efficacy outcome.
Results Of 1057 screened trials, 52 were identified as randomised trials with SF-36 results (66 separate comparisons). Only eight trials reported all 10 SF-36 scores in the published articles. For 21 of the 66 comparisons, SF-36 results were discordant for statistical significance compared with the results for primary efficacy outcomes. Of 17 statistically significant SF-36 scores, where primary outcomes were not also statistically significant in the same direction, the magnitude of effect was small in six, moderate in six, large in three, and not reported in two. Authors modified the interpretation of study findings based on SF-36 results in only two of the 21 discordant cases. Among 100 additional randomly selected trials not reporting any SF-36 information, at least five had collected SF-36 data but only one had analysed it.
Conclusions SF-36 measurements sometimes produce different results from those of the primary efficacy outcomes but rarely modify the overall interpretation of randomised trials. Quality of life and health related survey information should be utilised more systematically in randomised trials.

INTRODUCTION

Quality of life outcomes and surveys of health status are considered useful in randomised trials.[1-7] We evaluated published trials to determine whether it is common for quality of life and health survey results to reach different conclusions from those of the primary efficacy outcomes, whether there is selective reporting of outcomes, and whether discordant results in these outcomes modify the conclusions of trials.

METHODS

We considered randomised trials with data on SF-36 published in 2005 in five major general medicine journals and 17 high impact factor specialty journals

(see bmj.com). Trials were eligible if they reported on any of the two composite SF-36 scores (physical, mental) and eight subdomains (physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, mental health). When information was not reported on all 10 scores, we contacted the authors. We also considered trials using SF-12, a shorter version of SF-36 (for composite scores).

We screened articles identified through PubMed for keywords (see bmj.com). Suitable articles were evaluated independently by AK and IK. Disagreements were resolved by consensus or by DGC-I. To probe whether SF-36 data may have remained unpublished we emailed the authors of 100 randomly selected trials, among those retrieved trials not reporting SF-36 data. Data were extracted independently by IK, AK, and DGC-I, with discrepancies resolved by consensus or by JPAI.

From eligible articles we extracted information on authors, journal, design (superiority or non-inferiority), condition, interventions compared, sample size (randomised, analysed for SF-36), definition of primary efficacy outcome (as reported; if not clarified, we selected the outcome used for sample size calculations), time points and statistical analysis for the primary outcome and SF-36 assessments, whether SF-36 was a co-primary outcome, and whether any other quality of life and related health survey scales were used.

For the primary efficacy outcome and for each of the SF-36 assessments we recorded whether the difference between compared arms was statistically significant ($P<0.05$) favouring the experimental arm, non-statistically significant, or favoured the control arm. For trials with more than two arms we considered the comparison of each experimental intervention against control separately.

We considered SF-36 results as statistically significant when at least one of the composite or subdomain scores showed a statistically significant result in favour or against the experimental intervention.

Data on SF-36 outcomes were extracted for the reported analyses that corresponded as closely as possible to the time points as for primary outcomes. If the primary outcome was a time to event analysis or incorporated serial longitudinal measurements, we preferred the analysis of serial longitudinal SF-36 measurements; if this was unavailable, we recorded whether there was formal statistically significant difference at any time points when SF-36 had been appraised. When the primary outcome was appraised at a single time point, we recorded the SF-36 outcomes at the single same (or closest) time point. In two comparisons where co-primary outcomes existed and could not be prioritised, we based the evaluation of statistical significance on overall authors' interpretation.

For statistically significant SF-36 effects when the respective primary efficacy outcome was discordant, we extracted information on the effect size of SF-36 (see bmj.com). For comparisons with discordant statistical significance on SF-36 and primary outcome results, we recorded whether the authors had discussed the SF-36 results, whether they commented on the discrepancy and if so with what arguments, and if SF-36 findings changed the interpretation of the trial results.

RESULTS

Overall, 52 eligible trials with 66 comparisons were identified[w1-w52] (see bmj.com). Additional data were presented in other published articles on primary efficacy for one trial[w43] and SF-36 for eight trials.[w4 w21 w24 w29 w35 w36 w46 w51] Additional SF-36 data were provided by the authors in 11 trials (13 comparisons). Forty two trials (56 comparisons) addressed superiority, and 10 (10 comparisons) non-inferiority. In seven trials (10 comparisons)[w2 w8 w35 w39 w40 w44 w45] SF-36 was described as a co-primary outcome. Additional quality of life or health survey instruments appeared in 16 trials (16 comparisons).

Data for physical composite score and mental composite score were available for 34 trials (39 comparisons) and 35 trials (40 comparisons). Data on at least one of the eight subdomain scores were available for 36 trials (48 comparisons). Data on all possible SF-36 scores were available for 18 trials. Six trials[w6 w23 w29 w31 w35 w44] had collected information a priori only for specific subdomains.

Concordance of results

Of the 66 comparisons, 21 (32%) had discordant statistical significance for primary efficacy and SF-36 results (table). Moreover, of the 56 comparisons of superiority trials 19 had discordant results.

In one[w44] of the 21 discrepancies, SF-36 was a co-primary outcome. In seven discrepancies, additional health survey instruments were used. In two trials [w14 w51] these instruments agreed with SF-36, and in five[w12 w15 w31 w44 w47] trials with the primary efficacy outcome.

Among the 13 discordant comparisons with only SF-36 significant results, there were 17 statistically significant specific scores (five normalised, 10 raw, two reporting only statistical significance without effect size); effect sizes were small in six, moderate in six, and large in three.

Interpretation of trial findings in discordant settings

Improved primary outcome only—SF-36 results did not modify the trial's interpretation of these 11 comparisons (eight trials, see bmj.com).[w4 w12 w16 w18 w41-w43 w51] In five comparisons (four trials), SF-36 outcomes were only tabulated or alluded to in the results.[w12 w16 w18 w42] In the other four trials the authors focused on other non-primary outcomes,[w4] claimed that SF-36 was not sensitive enough to detect improvements,[w41] adopted a non-intention to treat analysis for SF-36 with significant results,[w43] or dismissed the importance of the negative effects on SF-36 in the face of benefits in disease-free survival.[w51]

Improved SF-36 only—SF-36 modified the interpretation of only two trials[w31 w44] (see bmj.com). In the other five comparisons (three trials), benefits on SF-36 did not change the interpretation.[w7 w14 w21] One trial dismissed the SF-36 difference as transient and weak,[w14] one considered the non-statistically significant benefits in efficacy as clinically important, whereas the significant improvements in SF-36 vitality scores were considered clinically unimportant and the authors questioned the use of SF-36 in trials on diabetes,[w21] and in another trial the authors considered that the clinical significance of statistically significant differences in SF-36 domains in patients with fibromyalgia could not be evaluated.[w7]

Improved SF-36, non-inferiority on primary outcome—SF-36 did not modify the interpretation of these two trials.[w46 w47] Both already concluded favourably for the experimental intervention that achieved the desired non-inferiority, and in one of them[w47] the observed benefit in SF-36 was considered due to chance.

Only SF-36 worsened—In one trial[w15] where SF-36 worsened with the experimental intervention, the investigators interpreted the results as showing no consistent differences in quality of life, because an additional instrument (EQ5D) showed no significant differences.

Probing unpublished data

Authors of 69 of 100 randomly selected trials responded. SF-36 data had been collected from five trials. The data had been analysed for only one trial and did not show any statistically significant differences for SF-36 or the primary efficacy outcome.

DISCUSSION

In one third of the trial comparisons in our empirical evaluation, differential effects on primary efficacy outcomes compared with SF-36 were identified. However, when SF-36 compared with efficacy outcomes reached discordant conclusions, SF-36 rarely affected the interpretation of these trials. What we observed was generally a tendency to belittle rather than to pronounce discordant results. Several trials did not discuss the SF-36 findings at all and most did not report all the tested SF-36 scores. Considering post hoc an instrument as insensitive or not worth reporting contradicts the initial choice to use this instrument as a trial outcome.

In most trials for chronic conditions, quality of life and surveys of health status are useful to consider. SF-36 was reported in fewer than 5% of the trials we screened and our author survey suggested that some additional trials (at least five of 100) had collected information on SF-36 but without analysing or publishing it two or three years after publication of the main trial results. Quality of life seems to remain undervalued in clinical research: few trials collected quality of life related data, fewer reported on them, data were only partially presented, and quality of life rarely affected the trial interpretation.

We should acknowledge some caveats. Firstly, by selecting high impact journals we identified trials with high visibility and probably high quality.[8] It is unlikely that this strategy would have selected for discordant results between outcomes. Secondly, selective analysis and reporting bias may affect primary outcomes and not just SF-36,[9-12] but this should not have increased the perceived rate of discrepancies between outcomes. Thirdly, discordance at the level of statistical significance does not necessarily mean that results for different outcomes differ beyond chance. Among statistically significant results, chance findings and non-clinically important differences are possible and primary outcomes should be given more weight in the discussion than secondary outcomes. Given that trials are typically powered to address the primary outcome, a significant result in the primary outcome with a non-significant result in quality of life or health survey assessments, may sometimes simply reflect lack of power for the quality of life or health survey outcome. Therefore we also examined the SF-36 effect sizes and the circumstances and discussion of discordant results. Fourthly, we did not carry out the same in-depth evaluation for trials where efficacy and SF-36 outcomes were concordant. It is unlikely that authors would then have modified their inferences, but SF-36 may have strengthened the conclusions. Finally, we did not examine trials using only other quality of life or health survey instruments beyond SF-36. However, SF-36 is the most

WHAT IS ALREADY KNOWN ON THIS TOPIC

Quality of life and related health survey outcomes could be essential in deciding whether an intervention is worth adopting

It is unknown whether such outcomes reach different conclusions from those of primary efficacy outcomes or whether they affect the interpretation of current clinical trials

WHAT THIS STUDY ADDS

Several randomised trials published in influential journals have had discordant results on primary efficacy outcomes compared with SF-36

When SF-36 and efficacy outcomes reached discordant conclusions, SF-36 rarely modified the interpretation of these trials

Concordance of statistical significance in SF-36 and primary outcome results

	SF-36 results			
Primary outcome	Significant*	Non-significant	Significant (against)†	Total
Significant	21	8	3	32
Non-significant	9‡	23	1	33
Significant (against)	0	0	1	1
Total	30	31	5	66

κ coefficient 0.33 (95% confidence interval 0.06 to 0.59) for concordance of primary outcome against SF-36. No situations occurred where specific SF-36 scores were significant for experimental intervention and others were significant against.
*At least one of composite or subdomain scores shows statistically significant result in favour of experimental intervention.
†At least one of composite or subdomain scores shows statistically significant result against experimental intervention.
‡This category contains the only two studies (w31 and w44) where interpretation of study findings was modified based on SF-36 results.

robustly standardised and widely used one and we wanted to maximise comparability.

Overall, quality of life and health survey assessments provide a different window into patient outcomes and deserve to be included in more trials with complete reporting of results, and standardised interpretation. Unbiased data on these outcomes may enhance our ability to improve clinical decision making.

We thank the following for clarifying their results or providing additional data: N Assefi, D Buchwald, and C Jacobsen (for Assefi et al[w1]); A Avenell and JA Cook (for Avenell et al[w2]); J Carratala, LJ Crofford, J Pepper, and B Lees (for De Arenaza et al[w10]); H Escobar-Morreale, RM Greenhalgh, and LC Brown (for EVAR 1[w14] and EVAR 2[w15]); I Gilron, S Hewlett, F Hill-Briggs, CA Hukins, P Jellema, and DA van der Windt (for Jellema et al[w23]); JA Klaber-Moffett, K Linde, PS Parfrey, and RN Foley (for Parfrey et al[w31]); DJ Torgerson (for Porthouse et al[w32]); D Revicki and JM Miranda (for Revicki et al[w35]); M Rienstra and IC van Gelder (for Rienstra et al[w36]); BL Rollman, MF Scheier, and S Colvin (for Scheier et al[w40]); N Shaheen, BN Singh, AL Stanton, and G Bleijenberg (for Stulemeijer et al[w45]); D van der Heijde and JH Stone (for Wegener's Granulomatosis Etanercept Trial[w50]); and WA Whitelaw.

Contributors: See bmj.com.

Funding: None.

Competing interests: None declared.

Ethical approval: Not required.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Guyatt GH, Naylor CD, Juniper E, Heyland DK, Jaeschke R, Cook DJ. Users' guides to the medical literature. XII. How to use articles about health-related quality of life. Evidence-Based Medicine Working Group. *JAMA* 1997;277:1232-7.

2 Sanders C, Egger M, Donovan J, Tallon D, Frankel S. Reporting on quality of life in randomised controlled trials: bibliographic study. *BMJ* 1998;317:1191-4.

3 Fayers PM, Hopwood P, Harvey A, Girling DJ, Machin D, Stephens R. Quality of life assessment in clinical trials—guidelines and a checklist for protocol writers: the UK Medical Research Council experience. MRC Cancer Trials Office. *Eur J Cancer* 1997;33:20-8.

4 Food and Drug Administration. Draft guidance for industry on patient-reported outcome measures: use in medicinal product development to support labeling claims. *Fed Register* 2006;71:5862-3.

5 Committee for Medicinal Products for Human Use. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. www.emea.europa.eu/pdfs/human/ewp/1393104en.pdf. 2005.

6 Revicki DA. Regulatory Issues and Patient-Reported Outcomes Task Force for the International Society for Quality of Life Research. FDA draft guidance and health-outcomes research. *Lancet* 2007;369:540-2.

7 Revicki DA, Gnanasakthy A, Weinfurt K. Documenting the rationale and psychometric characteristics of patient reported outcomes for labeling and promotional claims: the PRO evidence dossier. *Qual Life Res* 2007;16:717-23.

8 Nieminen P, Carpenter J, Rucker G, Schumacher M. The relationship between quality of research and citation frequency. *BMC Med Res Methodol* 2006;6:42-9.

9 Ioannidis JP, Trikalinos TA. An exploratory test for an excess of significant findings. *Clin Trials* 2007;4:245-53.

10 Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ* 2005;330:753.

11 Chan AW, Krleza-Jerić K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ* 2004;171:735-40.

12 Chan AW, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291:2457-65.

Accepted: 19 September 2008

Comorbidity and repeat admission to hospital for adverse drug reactions in older adults: retrospective cohort study

Min Zhang, C D'Arcy J Holman, Sylvie D Price, Frank M Sanfilippo, David B Preen, Max K Bulsara

EDITORIAL by Morgan

School of Population Health, University of Western Australia, Perth, WA, Australia

Correspondence to: M Zhang
min.zhang@uwa.edu.au

Cite this as: *BMJ* 2009;338:a2752
doi:10.1136/bmj.a2752

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;338:a2752

ABSTRACT

Objectives To identify factors that predict repeat admission to hospital for adverse drug reactions (ADRs) in older adults.

Design Population based retrospective cohort study.

Setting All public and private hospitals in Western Australia.

Participants 28 548 patients aged ≥60 years with an admission for an ADR during 1980-2000 followed for three years using the Western Australian data linkage system.

Results 5056 (17.7%) patients had a repeat admission for an ADR. Repeat ADRs were associated with sex (hazard ratio 1.08, 95% confidence interval 1.02 to 1.15, for men), first admission in 1995-9 (2.34, 2.00 to 2.73), length of hospital stay (1.11, 1.05 to 1.18, for stays ≥14 days), and Charlson comorbidity index (1.71, 1.46 to 1.99, for score ≥7); 60% of comorbidities were recorded and taken into account in analysis. In contrast, advancing age had no effect on repeat ADRs. Comorbid congestive cardiac failure (1.56, 1.43 to 1.71), peripheral vascular disease (1.27, 1.09 to 1.48), chronic pulmonary disease (1.61, 1.45 to 1.79), rheumatological disease (1.65, 1.41 to 1.92), mild liver disease (1.48, 1.05 to 2.07), moderate to severe liver disease (1.85, 1.18 to 2.92), moderate diabetes (1.18, 1.07 to 1.30), diabetes with chronic complications (1.91, 1.65 to 2.22), renal disease (1.93, 1.71 to 2.17), any malignancy including lymphoma and leukaemia (1.87, 1.68 to 2.09), and metastatic solid tumours (2.25, 1.92 to 2.64) were strong predictive factors. Comorbidities requiring continuing care predicted a reduced likelihood of repeat hospital admissions for ADRs (cerebrovascular disease 0.85, 0.73 to 0.98; dementia 0.62, 0.49 to 0.78; paraplegia 0.73, 0.59 to 0.89).

Conclusions Comorbidity, but not advancing age, predicts repeat admission for ADRs in older adults, especially those with comorbidities often managed in the community. Awareness of these predictors can help clinicians to identify which older adults are at greater risk of admission for ADRs and, therefore, who might benefit from closer monitoring.

INTRODUCTION

Adverse drug reactions (ADRs) are a major public health problem in older populations.[1-3] In western countries, ADRs cause 3-6% of all hospital

admissions[1-3] and are responsible for about 5-10% of inpatient costs.[4-6] In Western Australia by 2003 over 30% of all ADRs were repeat ADRs.[7]

Despite concerns that ADRs represent an important medical problem in older people, the predictive factors are still poorly understood. Risk factors independently associated with ADRs have included advancing age, sex, comorbidity, multiple drug regimens, inappropriate use of medication, alcohol intake, poor cognitive function, and depression.[4 8-13] There is no consensus on which factors have the greatest impact. This retrospective cohort study is based on a state-wide population of patients and investigates whether comorbid conditions, demographic factors, and drug category are associated with a repeat admission for ADRs in people aged ≥60.

METHODS

Study setting and population

We used administrative data from all public and private hospitals in Western Australia, a state with a population of 2.09 million in 2007.[14] The study population consisted of all residents aged ≥60 with a hospital admission related to an ADR identified through the data linkage system. In February 2005 we extracted linked hospital and death records for all patients aged ≥60 with an admission for ADR in 1980-2003.

Table 1 | Drug categories (from ICD external cause codes) responsible for admission for first time adverse drug reaction (ADR) in patients aged ≥60, Western Australia, 1980-2000

Drug category (as defined in ICD-10-AM)	No (%) of patients
Systemic antibiotics*	2672 (9.4)
Other systemic anti-infectives/antiparasitics	430 (1.5)
Hormones (including synthetics and antagonists)	2174 (7.6)
Primarily systemic agents	2122 (7.4)
Agents primarily affecting blood constituents	2432 (8.5)
Analgesics/antipyretics/anti-inflammatory drugs	4694 (16.4)
Antiepileptics/antiparkinsonian drugs	1121 (3.9)
Sedatives, hypnotics, antianxiety drugs	406 (1.4)
Anaesthetics and therapeutic gases	414 (1.5)
Psychotropic drugs†	1534 (5.4)
Central nervous system stimulants	18 (0.1)
Drugs primarily affecting autonomic nervous system‡	476 (1.7)
Agents primarily affecting cardiovascular system	5576 (19.5)
Agents primarily affecting gastrointestinal system	252 (0.9)
Agents affecting water balance/minerals/uric acid§	2024 (7.1)
Agents affecting muscles/respiratory system	563 (2.0)
Topical agents affecting skin, eyes, ENT, dental	577 (2.0)
Other and unspecified drugs and medicines	982 (3.4)
Bacterial vaccines	16 (0.1)
Other and unspecified vaccines/biologicals	65 (0.2)
Total	28 548 (100.0)

ADR=adverse drug reaction; ICD=International Classification of Diseases; ENT=ear, nose, throat.
*Excludes antineoplastic antibiotics (E930.7) from ICD-9/ICD-9-CM (were added to primarily systemic agents, which include antineoplastics).
†Excludes benzodiazepines (E939.4) from ICD-9/ICD-9-CM (were added to group sedatives, hypnotics, antianxiety drugs, which includes benzodiazepines in ICD-10).
‡Excludes sympatholytics (E941.3) from ICD-9/ICD-9-CM (were added to agents primarily affecting cardiovascular system, which include these in ICD-10).
§Excludes theophylline (E944.1) from ICD-9/ICD-9-CM (was added to agents affecting muscles/respiratory system, which includes antiasthmatics).

Definition of ADR and identification of patients

We included all ADRs that resulted in hospital admission or that occurred while patients were in hospital and extended the length of hospital stay. An ADR event was any adverse effect caused by correct use of drugs, medicines, or biological substances properly administered in therapeutic or prophylactic doses, excluding errors in the technique of administration of drugs, intentional and unintentional overdose, and abuse of a drug. An ADR hospital episode was defined as a period of continuous treatment for an ADR in one or more hospitals. We checked all records for transfers between hospitals and, where they existed, combined the admissions into a single "episode" for analysis.

Follow-up and outcome measurements

The length of follow-up was the time in days from hospital separation for the first time ADR episode (time zero) to the date of a second separate admission for an ADR or, in the absence of a repeat event within three years, to the date of the third anniversary from time zero or date of death if within three years (censored).

In the 3.9% of all instances when multiple drugs were thought to be responsible for an ADR we included in the analysis only the drug clinically considered to be primarily responsible. Drugs were grouped into 20 broad categories as defined in ICD-10-AM, with closest possible equivalent specifications for ICD-9 and ICD-9-CM. Table 1 shows the distributions of admissions for first time ADRs according to these drug categories.

We assigned a socioeconomic disadvantage score for each patient by transforming residential postcode into numerical values of the index of relative socioeconomic disadvantage, used extensively in public health research.[15]

Statistical analysis

Statistical analysis was performed with SPSS version 12.0.1 (SPSS, Chicago, IL). We used the independent samples t test for continuous variables and a χ^2 test for categorical variables to compare characteristics between patients with first time only and repeat ADR episodes. We obtained hazard ratios and associated 95% confidence intervals after adjusting for age, sex, indigenous status, residential locality, socioeconomic disadvantage, admission type, hospital type, length of hospital stay, calendar period of ADR, Charlson comorbidity index, and drug categories. Other studies have identified these variables as those influencing the risk of first time ADRs,[8-13] and they were significant predictors of repeat ADRs based on the preceding univariate analysis. To assess potential for survival bias, we separately analysed data from all patients and from a subset that excluded those who died in the three year follow-up period.

RESULTS

Results using data from all patients were similar to those obtained when we excluded patients who died in the follow-up period, so we report only the results from analyses that included all patients. Within the first three years of follow-up, 5056 patients (17.7%) experienced a

Table 2 | Adjusted hazard ratios (HR) with 95% confidence intervals for repeat adverse drug reactions (ADRs) for Charlson comorbidity at first admission to hospital for ADR in older adults

Comorbid conditions	No (%) with first time ADR only	No (%) with repeat ADR	Adjusted HR (95% CI)*	P value†
Congestive cardiac failure	3858 (16.4)	889 (17.6)	1.56 (1.43 to 1.71)	<0.001
Peripheral vascular disease	978 (4.2)	196 (3.9)	1.27 (1.09 to 1.48)	<0.01
Chronic pulmonary disease	1537 (6.5)	497 (9.8)	1.61 (1.45 to 1.79)	<0.001
Rheumatological disease	622 (2.6)	210 (4.2)	1.65 (1.41 to 1.92)	<0.001
Mild liver disease	164 (0.7)	35 (0.7)	1.48 (1.05 to 2.07)	0.03
Diabetes (mild to moderate)	2276 (9.7)	531 (10.5)	1.18 (1.07 to 1.30)	<0.01
Diabetes with chronic complications	507 (2.2)	248 (4.9)	1.91 (1.65 to 2.22)	<0.001
Renal disease	1045 (4.4)	373 (7.4)	1.93 (1.71 to 2.17)	<0.001
Any malignancy including lymphoma and leukaemia	2538 (10.8)	811 (16.0)	1.87 (1.68 to 2.09)	<0.001
Moderate or severe liver disease	89 (0.4)	19 (0.4)	1.85 (1.18 to 2.92)	<0.01
Metastatic solid tumour	1245 (5.3)	273 (5.4)	2.25 (1.92 to 2.64)	<0.001
Cerebrovascular disease	1601 (6.8)	226 (4.5)	0.85 (0.73 to 0.98)	0.02
Dementia	791 (3.4)	77 (1.5)	0.62 (0.49 to 0.78)	<0.001
Hemiplegia or paraplegia	762 (3.2)	102 (2.0)	0.73 (0.59 to 0.89)	<0.01
Myocardial infarction	1332 (5.7)	224 (4.4)	0.98 (0.84 to 1.13)	0.73
Peptic ulcer disease	1847 (7.9)	298 (5.9)	1.09 (0.95 to 1.25)	0.24
AIDS	3 (<0.1)	2 (<0.1)	1.65 (0.41 to 6.71)	0.48

*Estimates from Cox regression models included terms for age (continuous), gender (female, male), indigenous status (non-Aboriginal, Aboriginal), residence (regional/rural area, metropolitan), admission type (elective, emergency), type of hospital attended (public, private, other), length of stay (continuous), socioeconomic disadvantage (fifths), calendar period of ADR (continuous), and drug categories responsible for first-time ADR.
†For hazard ratio (HR), using patients with absent comorbidity as reference category.

repeat admission for an ADR. The adjusted hazard ratios for repeat ADRs were 1.08 (95% confidence interval 1.02 to 1.15) for men, 0.87 (0.80 to 0.93) for private hospital admissions, 1.11 (1.05 to 1.18) for length of hospital stay ≥ 14 days, 2.34 (2.00 to 2.73) for admission in 1995-9 compared with the earliest time period, and 1.71 (1.46 to 1.99) for a Charlson comorbidity index score ≥7, with a significant linear trend across quantitative or ordinal quantitative variables. Residential location in the metropolitan area (1.10, 1.01 to 1.19) was marginally significant (P=0.03). There was no significant effect on repeat ADRs for age, indigenous status, type of admission, and socioeconomic disadvantage.

Table 2 shows the effects of individual Charlson comorbid conditions on repeat ADRs. Compared with patients who had no recorded comorbidity, the analysis identified sizeable adjusted hazard ratios for congestive cardiac failure, peripheral vascular disease, chronic pulmonary disease, rheumatological disease, mild liver disease, mild to moderate diabetes, diabetes with chronic complications, renal disease, any malignancy including lymphoma and leukaemia, moderate to severe liver disease, and metastatic solid tumour. Cerebrovascular disease, dementia, and hemiplegia or paraplegia were apparently preventive for repeat ADRs. There was no significant relation for myocardial infarction, peptic ulcer disease, or AIDS, although people with AIDS had only two repeat ADRs.

DISCUSSION

Predictive factors for repeat ADR admission

We found strong evidence that comorbidity from chronic disease rather than advancing age increases rate of repeat ADRs in older adults. Comorbid congestive cardiac failure, diabetes, and peripheral vascular, chronic pulmonary, rheumatological, hepatic, renal, and malignant diseases were all strong predictors of readmissions for ADRs. We also found that comorbid cerebrovascular disease, dementia, and hemiplegia or paraplegia were associated with a reduced risk of repeat admission for ADRs. First admission for an ADR with a longer hospital stay, admissions in the most recent time period, and male sex also predicted repeat ADR admissions, with admission to private hospitals showing a reduced risk.

ADRs are acknowledged as a major health problem in older people.[1-7 16] A meta-analysis of 68 observational studies reported that the rates of admissions related to ADRs in older people were four times higher than in younger people.[17] However, we found that advancing age was not independently predictive of repeat admissions related to ADRs in people ≥60.

Strengths and limitations

There are several potential explanations for the observed importance of comorbidity. Comorbidity might increase vulnerability to ADRs by impairing body systems; there might be increased opportunity for drug interactions because of polypharmacy for multiple morbidities; and finally, Berkson's bias—that is, ADRs are more likely to be identified and diagnosed because of comorbid conditions increasing a person's contact with the health system.[18] A reduced risk of repeat ADRs associated with admission to private hospital might be explained by private patients having generally better health or being socioeconomically advantaged but the latter is inconsistent with the measure of least socioeconomic disadvantage used in

WHAT IS ALREADY KNOWN ON THIS TOPIC

Adverse drug reactions (ADRs) are a major public health problem in older populations

Repeat ADRs leading to hospital admission have increased at a greater rate than first time ADRs in older adults and by 2003 in Western Australia they had reached 30% of all ADRs

Little information is available on risk factors that predict repeat ADRs

WHAT THIS STUDY ADDS

Comorbid congestive cardiac failure, diabetes, peripheral vascular, chronic pulmonary, hepatic, renal, rheumatological, and malignant diseases predict readmission for ADRs

Comorbid cerebrovascular disease, dementia, and paraplegia seem to protect against repeat ADRs, possibly because such patients are under closer healthcare supervision

this study, which had no effect (hazard ratio 0.94, 0.87 to 1.02).

An important limitation of the study was the absence of data on drug doses and multiple drug regimens. We did have information on the drug category primarily responsible for ADRs and assessed the effect of each drug category by comparing it with the average risk of a repeat ADR from the 20 drug categories responsible for first time ADRs.

The strengths of our study include the cohort design with population based and audited data of high quality,[19] thus overcoming issues related to selection and recall biases as well as loss to follow-up. The longitudinal linked data allowed us to identify repeat ADRs in the same patient regardless of changes in the treating hospital. An important limitation of the study was that the administrative hospital morbidity data system is known to code only 60% of the 17 Charlson comorbid conditions relative to information obtained from chart review.[20] False positive diagnoses of comorbid conditions, however, are infrequent for most conditions (range 0-1.5%)[20] and it seems implausible that the levels of underascertainment would differ substantially between patients with first time ADR only and those with repeat ADRs.

As with other studies of this nature, reliability of ascertainment might vary because the presence and diagnosis of an ADR is subject to clinical judgment. The diagnosis of an ADR was made by senior hospital doctors (consultants and registrars) and junior doctors recorded them in text on a structured hospital separation abstract. The second step involved coding the text, including external causes or contributing factors. This was performed in each of the hospitals by qualified clinical coders who are trained in the use of the ICD codes and were able to obtain additional information from the medical notes as required. The accuracy of clinical coding (including E codes) is routinely checked by coding supervisors as well as by random internal audits.

We found that older adults who experienced an ADR during the most recent study period, 1995-9, had a 2.4-fold greater risk of recurrence than their counterparts in 1980-4. As the study was longitudinal, we need to consider the influence of factors that changed with time. The results are consistent with national drug consumption data that showed an increase in drug exposure in older Australians of 4.7% during 2000-1 alone.[21] This increase greatly exceeded population growth, suggesting either a larger population at risk or a higher average level of drug exposure per patient. Our results, derived from population level data, suggest that there exists a strong temporal correlation between repeat ADRs in older adults and greater use of drugs in the community generally.[16 22]

We thank the Data Linkage Branch, WA Department of Health, in particular, Diana Rosman, manager, for her assistance.

Contributors: See bmj.com.

Funding: MZ is supported by a postdoctoral fellowship from the National Health and Medical Research Council (Australia).

Competing interests: None declared.

Ethical approval: This work was approved by the human research ethics committee of the University of Western Australia.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA* 1998;279:1200-5.
2 Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *BMJ* 2004;329:15-9.
3 Roughead EE. The nature and extent of drug-related hospitalisations in Australia. *J Qual Clin Pract* 1999;19:19-22.
4 Onder G, Pedone C, Landi F, Cesari M, Della Vedova C, Bernabei R, et al. Adverse drug reactions as cause of hospital admissions: results from the Italian Group of Pharmacoepidemiology in the Elderly (GIFA). *J Am Geriatr Soc* 2002;50:1962-8.
5 Suh DC, Woodall BS, Shin SK, Hermes-De Santis ER. Clinical and economic impact of adverse drug reactions in hospitalized patients. *Ann Pharmacother* 2000;34:1373-9.
6 Moore N, Lecointre D, Noblet C, Mabille M. Frequency and cost of serious adverse drug reactions in a department of general medicine. *Br J Clin Pharmacol* 1998;45:301-8.
7 Zhang M, Holman CDJ, Preen DB, Brameld KJ. Repeat adverse drug reactions causing hospitalization in older Australians: a population-based longitudinal study 1980-2003. *Br J Clin Pharmacol* 2007;63:163-70.
8 Field TS, Gurwitz JH, Avorn J, McCormick D, Jain S, Eckler M, et al. Risk factors for adverse drug events among nursing home residents. *Arch Intern Med* 2001;161:1629-34.
9 Carbonin P, Pahor M, Bernabei R, Sgadari A. Is age an independent risk factor of adverse drug reactions in hospitalized medical patients? *J Am Geriatr Soc* 1991;39:1093-9.
10 Klarin I, Wimo A, Fastbom J. The association of inappropriate drug use with hospitalisation and mortality: a population-based study of the very old. *JAMA* 2005;293:2131-40.
11 Fialová D, Topinková E, Gambassi G, Finne-Soveri H, Jónsson PV, Carpenter I, et al. Potentially inappropriate medication use among elderly home care patients in Europe. *JAMA* 2005;293:1348-58.
12 Onder G, Landi F, Della Vedova C, Atkinson H, Pedone C, Cesari M, et al. Moderate alcohol consumption and adverse drug reactions among older adults. *Pharmacoepidemiol Drug Saf* 2002;11:385-92.
13 Onder G, Penninx BW, Landi F, Atkinson H, Cesari M, Bernabei R, et al. Depression and adverse drug reactions among hospitalized older adults. *Arch Intern Med* 2003;163:301-5.
14 Australian Bureau of Statistics. *Australian demographic statistics*, Mar 2007. Canberra: Australian Bureau of Statistics, 2007. (Report No. 3101.0.)
15 Australian Bureau of Statistics. *Census of population and housing: socio-economic indexes for areas*. Canberra: ABS, 2006. (Catalogue No: 2912.0, 2012.0, 2039.0.)
16 Burgess CL, Holman CD, Satti AG. Adverse drug reactions in older Australians, 1981-2002. *Med J Aust* 2005;182:267-70.
17 Beijer HJ, de Blaey CJ. Hospitalisations caused by adverse drug reactions (ADR): a meta-analysis of observational studies. *Pharm World Sci* 2002;24:46-54.
18 Berkson J. Limitations of the application of fourfold table analysis to hospital data. *Biometrics Bull* 1946;2:47-53.
19 Holman CDJ, Bass J, Rouse IL, Hobbs MST. Population-based linkage of health records in Western Australia: development of a health services research linked database. *Aust N Z J Public Health* 1999;23:453-9.
20 Preen DB, Holman CD, Lawrence DM, Baynham NJ, Semmens JB. Hospital chart review provided more accurate comorbidity information than data from a general practitioner survey or an administrative database. *J Clin Epidemiol* 2004;57:1295-304.
21 Commonwealth Department of Health and Ageing. *Cost to government of pharmaceutical benefits*. Canberra: Health Access and Financing Division, Department of Health and Ageing, 2003.
22 Safety and Quality Council. *Second national report on patient safety, improving medication safety*. Canberra: Australian Council for Safety and Quality in Health Care, 2002.

Accepted: 1 October 2008