Making sense of big healthcare data – analysing longevity

Lisanne Gitsels, UEA

Content

- Healthcare databases:
  - Rise of healthcare databases
  - Strenghts & liminations
  - Designing cohort studies
- Case study:
  - Cohort selection using THIN database
  - Model specifications
  - Results & recommendations
- Overall conclusions

Healtcare databases in the UK

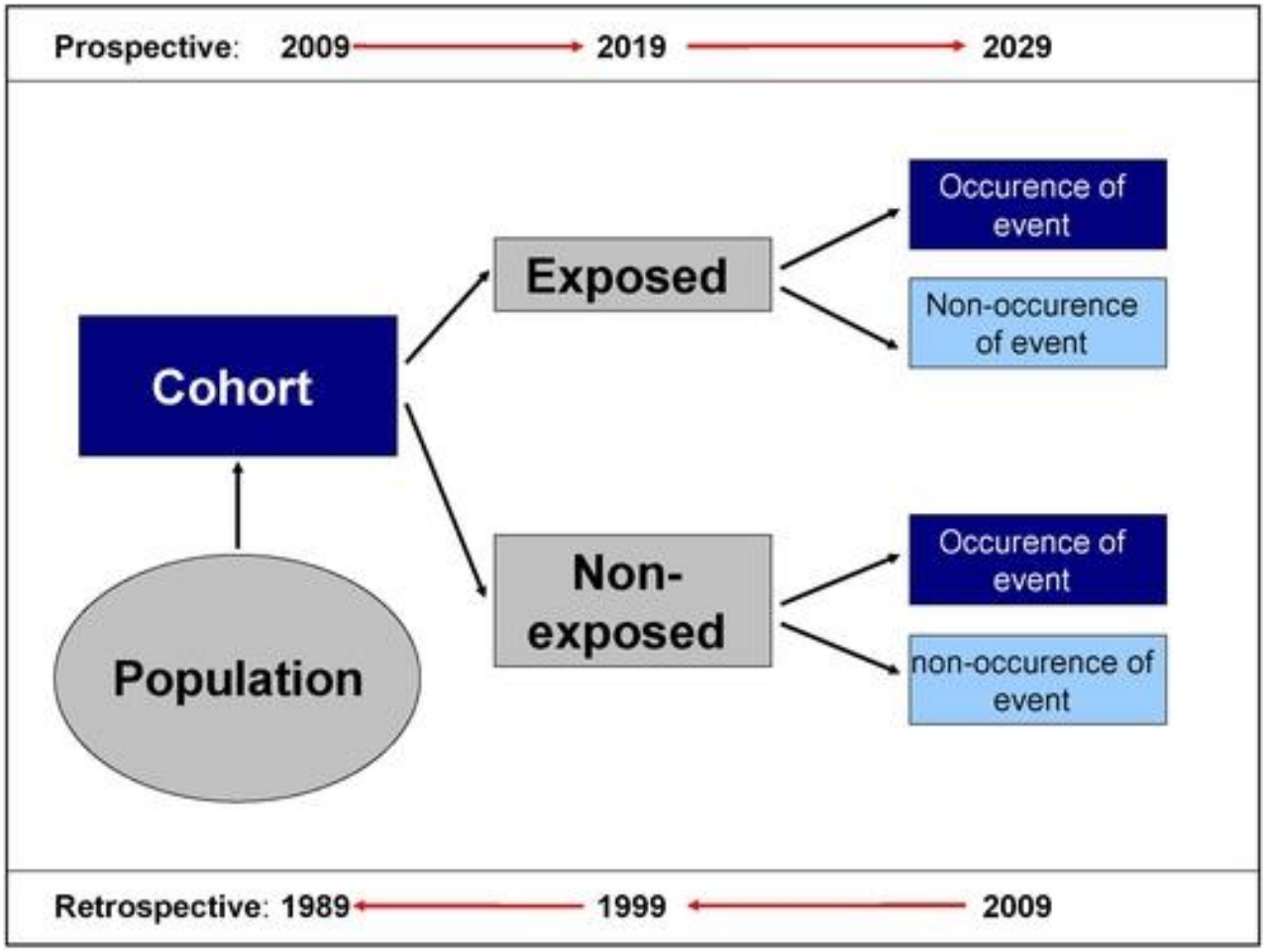
- Primary care:
  - GPRD/CPRD (General/Clinical Practice Research Datalink), since 1987
  - THIN (The Health Improvement Network), since 1987
  - QResearch, since 1993
- Secondary care:
  - HES (Hospital Episode Statistics): admitted patients since 1989, outpatients since 2003, and accidents & emergency records since 2007
- Emergency care:
  - ECDS (Emergency Care Data Set), since 2017

Strengths & limitations

+ Representative of UK
+ Reflects what happens in practice
+ Continually updated
+ Long follow-up
+ Low cost
+ Allow most epidemiological study designs
+ Linkage accross databases

- Missing or sporadic entries (e.g. blood pressure)
- No information on particular factor of interest (e.g. over-the-counter drugs)
- Bias by indication (healty user bias and sick user bias)

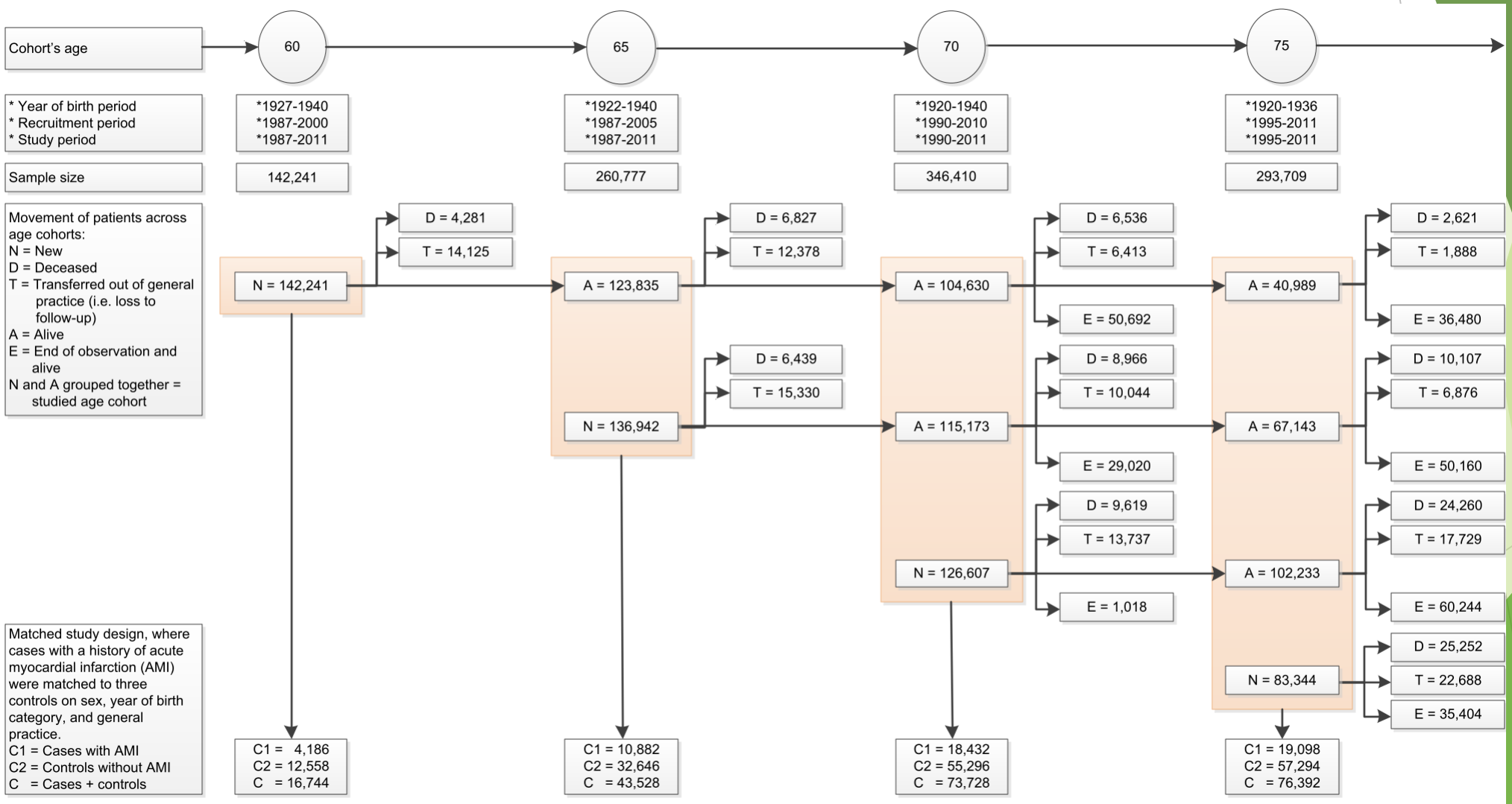
Designing cohort studies based on routine data



Case study

- Objectives: estimate the survival prospects associated with a history of a single or multiple heart attacks in the general population and estimate how the survival prospects were modified by recommended treatment.
- Gitsels LA, Kulinskaya E, Steel N Survival prospects after acute myocardial infarction in the UK: a matched cohort study 1987–2011 BMJ Open 2017;7:e013570. doi:10.1136/bmjopen-2016-013570.
- UEA's press release statement: https://www.uea.ac.uk/about/-/beta-blockers-offer-best-chance-of-increased-heart-attack-survival

Study design



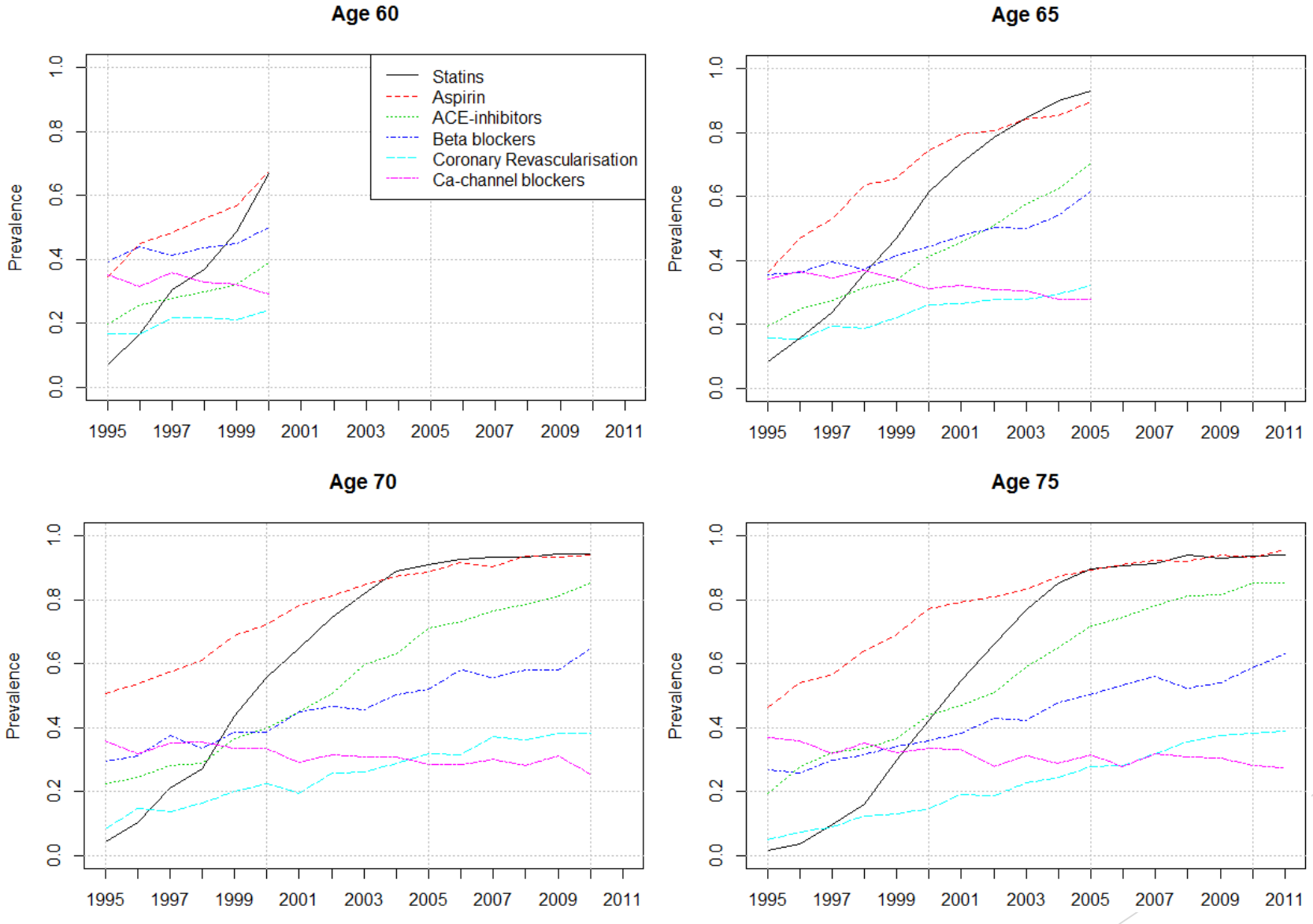
Data selection

- Outcome: time to death
- Primary exposure: heart attack (acute myocardial infarction)
- Treatments: coronary revascularisation (coronary artery bypass graft and coronary angioplasty), and prescription of ACE inhibitors, aspirin, beta blockers, calcium-channel blockers, and statins
- Confounders: sex, year of birth, socioeconomic status, angina, heart failure, other cardiovascular conditions (valvular heart disease, peripheral vascular disease, and cerebrovascular disease), chronic kidney disease, diabetes, hypertension, hypercholesterolaemia, alcohol consumption, body mass index, and smoking status
- Missing data dealt with by multiple imputation

Model specification

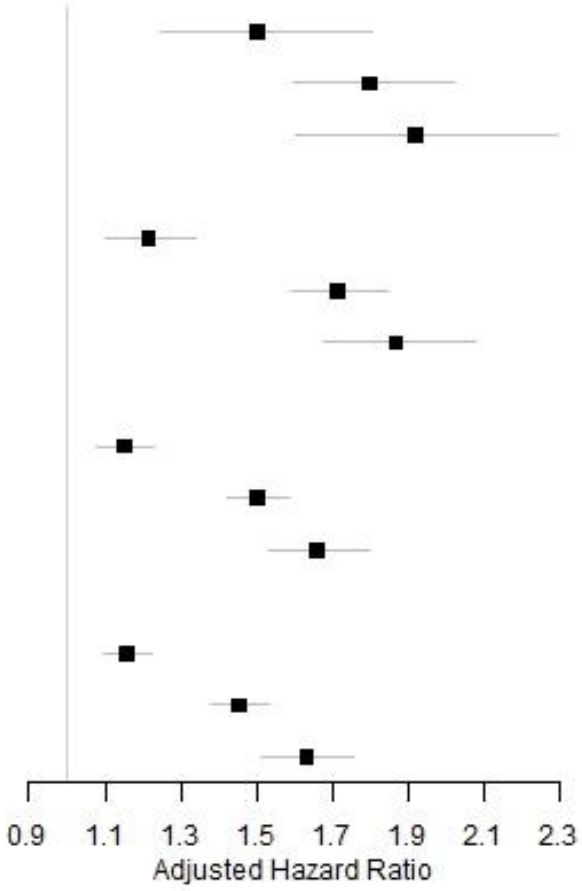
- Cox's proportional hazards regression estimates the hazard λ_{ij} for patient *i* from general practice *j*: $\lambda_{ij} = \lambda_0(t)\, Z_j\, e^{\beta X_{ij}}$
  - where λ_0 = baseline hazard (function of time),
  - Z_j = shared frailty term on general practice (constant),
  - β = coefficients (constant or time-variant),
  - and X_{ij} = exposures, e.g. heart attack (constant).
- Number of years gained or lost = β / ln (annual hazard of mortality)

Prevalence treatments in AMI patients



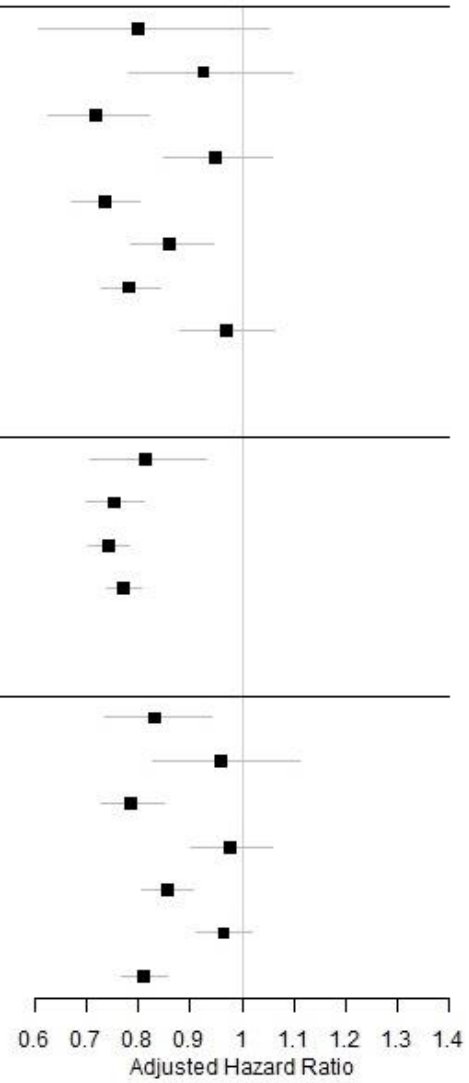
Survival prospects after AMI

Cohort	Ischaemic Heart Disease	Deaths (%per annum)	Adjusted HR (95%CI)
Age 60	No	1,843 (1.28)	
	Angina	165 (2.52)	1.50 (1.25-1.80)
	Single AMI	996 (2.56)	1.80 (1.60-2.02)
	Multiple AMIs	224 (2.89)	1.92 (1.60-2.29)
Age 65	No	5,180 (1.86)	
	Angina	602 (2.83)	1.21 (1.10-1.34)
	Single AMI	2,428 (3.19)	1.71 (1.59-1.84)
	Multiple AMIs	642 (3.81)	1.87 (1.68-2.07)
Age 70	No	9,264 (2.77)	
	Angina	1,293 (3.66)	1.15 (1.08-1.23)
	Single AMI	4,098 (4.38)	1.50 (1.42-1.59)
	Multiple AMIs	1,088 (5.14)	1.66 (1.53-1.80)
Age 75	No	10,686 (3.98)	
	Angina	1,988 (5.28)	1.16 (1.10-1.22)
	Single AMI	4,614 (6.02)	1.45 (1.38-1.53)
	Multiple AMIs	1,281 (7.22)	1.63 (1.51-1.76)



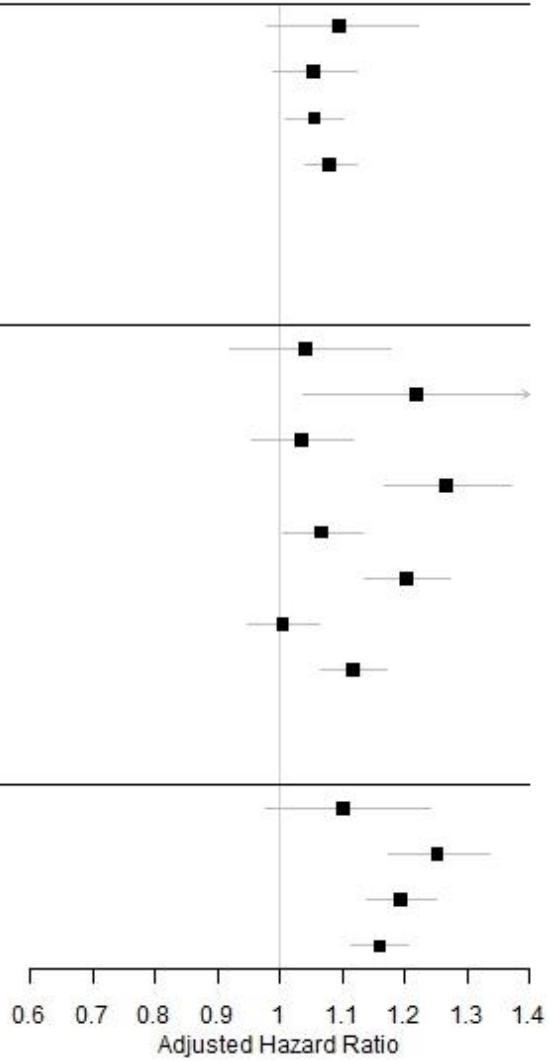
Survival prospects by treatments

Cohort	Coronary Revascularisation	Adjusted HR (95%CI)
Age 60	Follow-up<5yrs	0.80 (0.61-1.05)
	Follow-up>=5yrs	0.92 (0.78-1.10)
Age 65	Follow-up<5yrs	0.72 (0.63-0.82)
	Follow-up>=5yrs	0.95 (0.85-1.06)
Age 70	Follow-up<5yrs	0.73 (0.67-0.80)
	Follow-up>=5yrs	0.86 (0.78-0.94)
Age 75	Follow-up<5yrs	0.78 (0.73-0.84)
	Follow-up>=5yrs	0.97 (0.88-1.06)
	Statins	
Age 60	Yes	0.81 (0.71-0.93)
Age 65	Yes	0.75 (0.70-0.81)
Age 70	Yes	0.74 (0.70-0.78)
Age 75	Yes	0.77 (0.74-0.81)
	Beta blockers	
Age 60	Yes with AMI	0.83 (0.73-0.94)
	Yes without AMI	0.96 (0.83-1.11)
Age 65	Yes with AMI	0.79 (0.73-0.85)
	Yes without AMI	0.98 (0.90-1.06)
Age 70	Yes with AMI	0.85 (0.81-0.91)
	Yes without AMI	0.96 (0.91-1.02)
Age 75	Yes with AMI	0.81 (0.77-0.86)



Survival prospects by treatments (cont.)

Cohort	Aspirin	Adjusted HR (95%CI)
Age 60	Yes	1.10 (0.98-1.22)
Age 65	Yes	1.05 (0.99-1.12)
Age 70	Yes	1.05 (1.01-1.10)
Age 75	Yes	1.08 (1.04-1.12)
	Ca-channel blocker	
Age 60	Yes with AMI	1.04 (0.92-1.18)
	Yes without AMI	1.22 (1.04-1.43)
Age 65	Yes with AMI	1.03 (0.96-1.12)
	Yes without AMI	1.27 (1.17-1.37)
Age 70	Yes with AMI	1.07 (1.00-1.13)
	Yes without AMI	1.20 (1.14-1.27)
Age 75	Yes with AMI	1.00 (0.95-1.06)
	Yes without AMI	1.12 (1.06-1.17)
	ACE-inhibitor	
Age 60	Yes	1.10 (0.98-1.24)
Age 65	Yes	1.25 (1.17-1.34)
Age 70	Yes	1.19 (1.14-1.25)
Age 75	Yes	1.16 (1.11-1.21)



Recommendations

- Heart attack survivors are to a lesser extent worse off than previously estimated
- Survival benefits associated with coronary revascularisation and prescription of statins and beta blockers → more prescriptions
- Survival harms associated with prescription of aspirin and ACE inhibitors → further research
- Advocating equality in treatment

Overall conclusions

- Healthcare databases are beneficial to research
  - Insights in what happens in practice
- Healthcare databases allow for most epidemiological studies
  - E.g. design cohort study on longevity

Any questions or remarks ☺?