Imperial College London

Faculty of Medicine

Year 1 – 2012/13
Epidemiology in Practice
Autumn Term Course Guide

The Broad Street Pump, *Safe & Sound*, Penguin, 1971 **in** English MP. *Victorian Values -The Life and Times of Dr. Edwin Lankester*, 1990 http://www.ph.ucla.edu/epi/snow.html

Course Leader
Dr Mireille B. Toledano
Tel: 020 7594 3298
Email: m.toledano@imperial.ac.uk
Website: https://education.med.imperial.ac.uk

CONTENTS

SOLE FEEDBACK – Epidemiology in Practice

The following pages provide you with templates on which you can record your thoughts as the course proceeds. At the end of the course you can enter your views onto SOLE.

Please answer all questions by selecting the response which best reflects your view.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
The support materials available for this module (e.g. handouts, web pages, problem sheets and/or notes on the board).	□	□	□	□	□
The organisation of the module.	□	□	□	□	□
Feedback on my work has been prompt (this refers to your work being commented upon within a specified time).	□	□	□	□	□
Feedback on my work has helped me clarify things I did not understand.	□	□	□	□	□

Please use this box for constructive feedback and suggestions for improvement.

SOLE FEEDBACK – INDIVIDUAL LECTURERS

Please note that for SOLE, a Lecturer's name will only appear once. This template gives you the opportunity to record your comments about each lecture in the order of delivery.

On the following section, you have an opportunity to record any comments and constructive feedback you have for each lecturer.

	The lecture(s) are well structured					**The lecturer explains concepts clearly**					**The lecturer engages well with the students**				
Lecturer and Lecture Title	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Professor Sir Roy Anderson Global Health – Infectious Disease	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Dr Sarah Fidler AIDS – History And Progression	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Professor Paul Elliott Global Health – Non-Infectious Disease (Cardio)	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Professor Majid Ezzati Global Health – Non-Infectious Disease (Cancer)	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Dr Paul Aylin Why Evidence Based Medicine	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□

	The lecture(s) are well structured					The lecturer explains concepts clearly					The lecturer engages well with the students				
Lecturer and Lecture Title	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Dr Paul Aylin Association and Causation	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Dr Paul Aylin Video and Paper by Wakefield on MMR vaccine	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Dr Alex Bottle Descriptive Studies and Routine Data	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Dr Petra Wark Cohort and Case Control Studies	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Professor Helen Ward Clinical Trials	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Dr Tania Misra Public Health and Health Promotions Interventions	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□

	The lecture(s) are well structured					The lecturer explains concepts clearly					The lecturer engages well with the students				
Lecturer and Lecture Title	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Dr Teresa Norat Systematic Reviews and Meta Analysis	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Dr Claire Robertson Introduction to Critical Appraisal	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Dr. Susan Hodgson Understanding and appraising evidence	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Mr Mike Rowson Poverty, Health and Development/ Globalisation and Health Worker Migration	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Professor Alan Fenwick Waterborne Infectious Diseases in Africa	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□

	The lecture(s) are well structured					The lecturer explains concepts clearly					The lecturer engages well with the students				
Lecturer and Lecture Title	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Dr. Bhargavi Rao Disease Prevention – Screening	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Dr Mireille Toledano Strategic revision for exams	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□

MAPPING OF THEMES BY COURSE THROUGHOUT THE MBBS CURRICULUM

THEMES	***Global Health infectious disease***	***Global Health non-infectious disease***	***Evidence based medicine***	***Searching for evidence and creation of reference databases***	***Critical appraisal of evidence***	***Understanding bio-statistics***	***Public Health promotion and interventions***	***Screening, diagnostic tests and differential diagnosis***
Year 1	EIP, Society and Health	EIP, Society and Health	EIP	EIP, Library session: Managing your information: databases	EIP, 3rd PBL case	EIP	EIP	EIP, 3rd PBL case
Year 2	Science and Patient (theme called "sepsis"), PBL case 3		Science and Patient PBL – all four cases	Science and Patient PBL – all four cases, not databases	Science and Patient PBL – all cases, especially 3	Science and Patient	PBL case 1	Science and patient PBL case 2
Year 3								
Year 4	BSc Global Health	BSc Global Health					BSc Global Health	
Year 5								
Year 6								

INTRODUCTION

The ***Epidemiology in Practice*** course is taught in the Autumn Term of year 1.

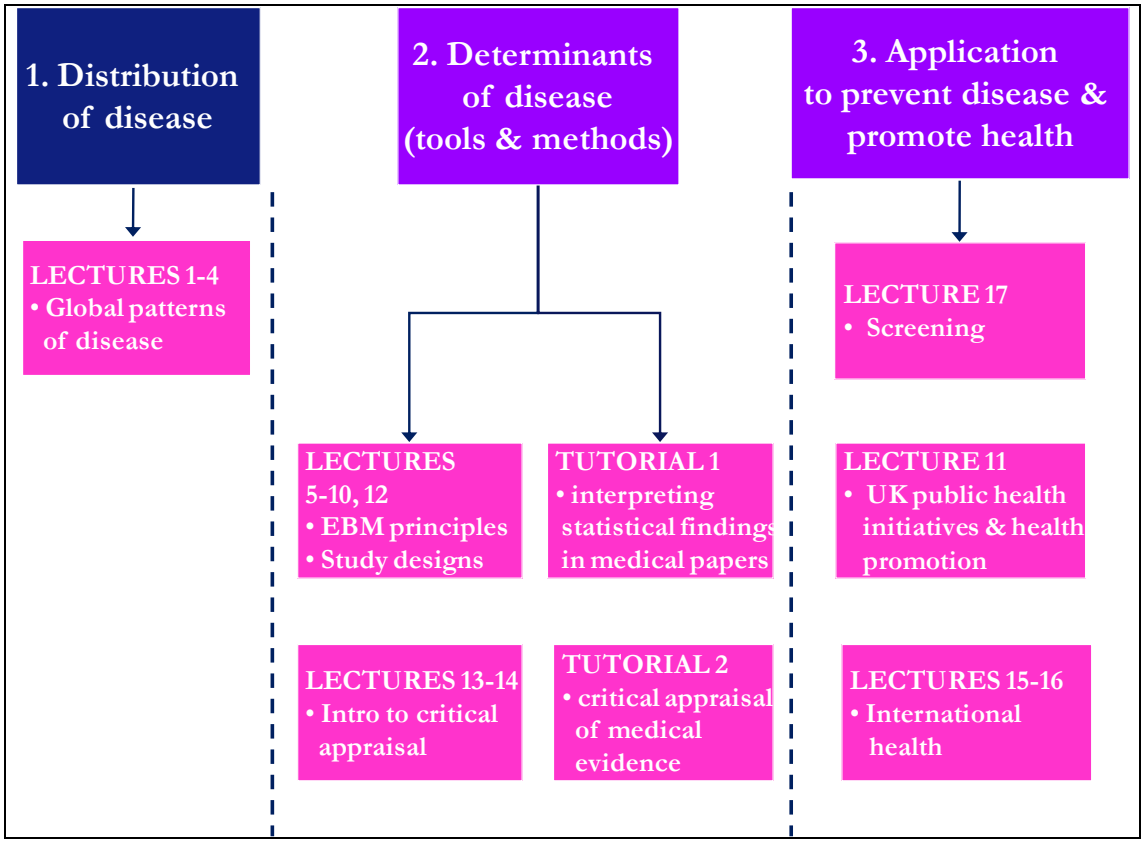
What is Epidemiology?

> "The study of the **distribution and determinants** of health related states or events in specified populations and the **application of this study to control health problems** - to promote, protect and restore health" (John Last, 1997)

COURSE STRUCTURE

The course consists of ten sessions. Each session lasts 3 hours. There are 18 lectures and 2 tutorials.

There are 3 core components to the course: 1. distribution/patterns of disease; 2. determinants of disease; 3. application to prevent disease and promote health (see flow chart of course structure below).



ASSESSMENT DETAILS

Formative Assessments

There will be a self test for this course.

Students will be able to take the self test once during the week 7th - 11th January 2013 at any convenient time. The self test will consist of EMQs and SBAs. Total recommended time is 55 minutes. This will be a good opportunity for you to practice exam questions that will be in exactly the same format as the summative exam. The exam will be accessible on the computers in SAFB using the LAPT (Blackboard Learn) system. Immediate marking of each question will operate. Students can look in advance on Blackboard and try exercises to see how it works.

Summative Assessment

As part of the Foundations of Clinical Practice (FoCP) theme, the course material for EIP, together with the Society and Health course, will be examined through EMQs and SBAs in a summative in-class test under examination conditions. For latest details of the format and dates of all examinations please refer to the Exams and Assessment page on the teaching intranet.

Examples of specimen questions

There are no specimen questions but the final revision lecture (lecture 18) will go through examples of the styles of questions likely to be used and some of these will be put on the intranet following the lecture. *Students are strongly advised to attend this final revision session*.

READING LIST

All examinable material will be presented in the lectures and tutorial classes. However, further reading always helps and it is strongly recommended! It will allow a better and more thorough understanding of the subject.

Core Course reading:

Ward H, Toledano M.B, Shaddick G, Davies B, Elliott P*, Oxford Handbook of* Epidemiology *for Clinicians (2012),* Oxford University Press

Recommended:

Bailey L, Vardulaki K, Langham J, Chandramohan D. (2005) *Introduction to Epidemiology*, Open University Press

Coggon D, Barker D, Rose G (2003) *Epidemiology for the uninitiated,* 5th edition, BMJ publishing.

Greenhalgh T. (2010) *How to read a paper: the basics of evidence-based medicine*, 4th Ed. Blackwell.

Bonita R, Beaglehole R, Kjellstrom T (2006) Basic Epidemiology, 2nd ed. WHO publishing.

Carr S, Unwin N, Pless-Mulloli T (2007) An introduction to public health and epidemiology. 2nd ed. OUP McGraw Hill.

Goldacre, B. (2009) Bad Science. Harper Perennial.

All course material (lecture slides and tutorial exercises with solutions) will be available on the college intranet. Log on to http://education.med.imperial.ac.uk – click on the Year 1 link, followed by the Foundations of Clinical Practice link, then the Epidemiology in Practice link.

TIMETABLE

Course sessions take place on both the South Kensington (SK) and Charing Cross (CX) campuses.

Details are correct at the time of going to press. Any amendments will be shown on the Course Timetable shown on the Intranet. Please check regularly during the course.

Date and campus	Time	Lecture topic	Lecturer
Session 1 Tuesday **(06/11/12)** **2–5 pm** SAFB–G16 Lecture Theatre	Lecture 1 2.00–3.30 pm	**Plenary Session: Global Patterns of Disease I** Global Health – infectious disease	Professor Sir Roy Anderson
	3.30–3.45 pm	**BREAK (15 minutes)**	
	Lecture 2 3.45–4.45pm	AIDS – history and progression (video and lecture)	Dr. Sarah Fidler
Session 2 Wednesday **(07/11/12)** 9.00am – 12.00pm Drewe Lecture Theatre CX	Lecture 3 9.30-10.45am	**Plenary Session: Global patterns of Disease II** Global Health – Non Infectious Disease (Cardio)	Professor Paul Elliott
	10.45– 11.00am	**BREAK (15 minutes)**	
	Lecture 4 11.00am–12.00 noon	Global Health – Non Infectious Disease (Cancer)	Professor Majid Ezzati
Session 3 Tuesday **(13/11/12)** 2.00 – 5.00pm Drewe Lecture Theatre CX	Lecture 5 2.00 – 3.00 pm	**The importance of evidence in the practice of medicine** Why evidence-based medicine?	Dr Paul Aylin
	Lecture 6 3.00 – 4.00 pm	Association and causation (covering sampling, chance – including introduction to p-values and confidence intervals -, bias, confounding, causal relationships)	Dr Paul Aylin

	Lecture 7 4.00-5.00 pm	Video and paper by Wakefield on MMR vaccine (as an example of what happens if you don't practice EBM!)	Dr Paul Aylin
Session 4 Wednesday **(14/11/12)** 9.00am – 12.00pm Drewe Lecture Theatre CX	Lecture 8 9.00-10.00am	**Study Design** Descriptive studies and routine data	Dr Alex Bottle
	10.00– 10.15am	**BREAK (15 Minutes)**	
	Lecture 9 10.15-11.15am	Cohort and case control studies	Dr Petra Wark
Session 5 Tuesday **(20/11/12)** 9.00am- 12.00pm SAFB-G16 Lecture Theatre SK	Lecture 10 9.00 - 10.00am	**More on Study Design** Clinical trials	Professor Helen Ward
	10.00– 10.15am	**BREAK (15 Minutes)**	
	Lecture 11 10.15 am – 11.15am	**Disease Prevention I** Public Health and Health Promotion Interventions	Dr. Tania Misra
Session 6 Tuesday **(20/11/12)** 2.00 – 5.00pm SAFB – MDL1 (Bay D) SAFB – MDL 1 & 2 (Bays A–C) SK	Tutorial 1 Group 1 2.00 – 3.30pm Group 2 3.30 – 5.00pm	**Tools of the trade (understanding and interpreting the statistical findings commonly reported in papers)** Group 1 (half of year, split into 7 groups of 20 students. Group 2 (half of year, split into 7 groups of 20 students). See page 6 for groups and venues	Tutors see page 6
Session 7 Thursday **(22/11/12)** 2.00 - 5.00pm SAFB-G16	Lecture 12 2.00 – 3.00 pm	**More on Study Design** Systematic reviews and meta analysis	Dr Teresa Norat
	3.00 – 3.15 pm	**BREAK (15 Minutes)**	

Lecture Theatre SAFB-MDL 1 (Bay D) SK	Lecture 13 3.15 – 4.15 pm	**Critical appraisal skills** Introduction to critical appraisal and how to read published papers	Dr Claire Robertson
Session 8 Tuesday 1.30 – 5.30pm **(04/12/12)** SAFB-G16 Lecture Theatre	Lecture 14 1.30 – 2.30 pm	**Understanding and appraising evidence** Horizon video and questions	Dr. Susan Hodgson
SAFB-MDL1 (Bays A – D) SAFB-MDL2 (Bays A – D) SK	Tutorial 2 2.30 – 4.00pm 4.00 – 5.30pm	**Critical appraisal of medical evidence** Group 1 (half of year, split into 7 groups of 20 students) Group 2 (half of year, split into 7 groups of 20 students) See page 6 for groups and venues	Tutors see page 6
Session 9 Wednesday 9am – 12pm **(5/12/10)** SAFB-G16 Lecture Theatre SK	Lecture 15 9.00 – 11.00am 11.00 – 11.30am Lecture 16 11.30am –- 12.30pm.	**International Health l Part 1** Poverty, Health and Development **International Health l Part 2** Globalisation and Health Worker Migration **BREAK** **International Health ll** The integration of global health partnerships to combat the propagation of waterborne infectious diseases in Africa	Mr Mike Rowson Professor Alan Fenwick
Session 10 Thursday 2pm – 5pm **(6/12/12)** SAFB-G16 Lecture Theatre SK	Lecture 17 2.00 – 3.00	**Disease Prevention ll** Screening	Dr. Bhargavi Rao
	3.00 – 3.15pm	**BREAK (15 Minutes)**	
	Lecture 18 3.15 – 5.00pm	**Strategic revision for exams**	Dr Mireille Toledano

TUTORIALS – GROUPINGS AND VENUES

Tutorial 1

For the tutorial, the year is split into two halves: first half 2 pm – 3.30 pm, second half 3.30 pm – 5 pm. There are roughly 10-20 students per group, as follows:

Group 1: Session 6, Tuesday 20th November 2012, 2.00 pm – 3.30 pm

Student Groups	A1/C1	A2/C2	A3/C3	A4/C4	A5/C5	A6/C6	A7/C7
Venue	MDL1A	MDL1B	MDL1C	MDL1 D	MDL2A	MDL2B	MDL2C
Tutor	Alex Bottle	Joanna Murray	Luciana Rubinstein	Minttu Ronn	Paul O'Reilly	Marc Chadeau-Hyam	Claudia Schoenborn

Group 2: Session 6, Tuesday 20th November 2012, 3.30 pm – 5.00 pm

Student Groups	B1/D1	B2/D2	B3/D3	B4/D4	B5/D5	B6/D6	B7/D7
Venue	MDL1A	MDL1B	MDL1C	MDL1 D	MDL2A	MDL2B	MDL2C
Tutor	Alex Bottle	Joanna Murray	Luciana Rubinstein	Minttu Ronn	Paul O'Reilly	Marc Chadeau-Hyam	Claudia Schoenborn

Tutorial 2

For the tutorial, the year is split into two halves: first half 2.30 pm – 4.00 pm, second half 4.00 pm – 5.30 pm. There are roughly 10-20 students per group, as follows:

Group 1: Session 8, Tuesday 4th December 2012, 2.30 pm – 4.00 pm

Student Groups	B1/D1	B2/D2	B3/D3	B4/D4	B5/D5	B6/D6	B7/D7
Venue	MDL1A	MDL1B	MDL1C	MDL1D	MDL2A	MDL2B	MDL2C
Tutor	Alex Bottle	Rachel Kelly	Petra Peeters	Vanessa Garcia Larson	Pauline Scheelbeek	Susan Hodgson	Rachel Smith

Group 2: Session 8, Tuesday 4th December 2012, 4.00 pm – 5.30 pm

Student Groups	A1/C1	A2/C2	A3/C3	A4/C4	A5/C5	A6/C6	A7/C7
Venue	MDL1A	MDL1B	MDL1C	MDL1D	MDL2A	MDL2B	MDL2C
Tutor	Alex Bottle	Rachel Kelly	Petra Peeters	Vanessa Garcia Larson	Pauline Scheelbeek	Susan Hodgson	Rachel Smith

LEARNING OUTCOMES

These course and session objectives provide you with a way to assess how well you are keeping up with the material. Note that they are also provided to the external examiners as a guide to what you should know at the end of the course.

General course learning outcomes

1. To describe global patterns of infectious and non-infectious disease, appreciate the disparities worldwide, and identify broad underlying causes for these patterns.
2. To appreciate the hierarchy of evidence in study design through knowledge of the strengths and weaknesses of various study designs, and to understand the importance of applying evidence to clinical decision making
3. To be able to understand and interpret the statistical findings commonly reported in scientific papers
4. To list and understand the main principles regarding how to read and critically evaluate a scientific paper
5. To describe, and give examples, of the main methods of intervention to improve health, on a national and international scale, including education, protection and prevention.

Session-specific learning outcomes

A. Sessions 1 & 2

- Define and distinguish incidence, prevalence, and mortality
- Describe the current burden of infectious diseases and their disparities worldwide.
- Identify the six commonest infectious causes of world mortality and some of the causes underlying their high incidence.
- Explain the concept of epidemiological transition.
- Describe the current burden of non-infectious diseases and their disparities worldwide.
- Identify the commonest non-infectious causes of world mortality and some of the causes underlying their high incidence.

B. Session 3

- Recognise the role of evidence based practice in clinical medicine
- List and define possible explanations for observed associations (chance, bias, confounding, causation), and cite examples of each
- Be able to describe the hierarchy of evidence in study design
- List the Bradford-Hill criteria for establishing causation and apply these to specific examples
- Be able to apply epidemiological skills to clinical decision making

C. Sessions 4, 5 (lecture 10), and 7 (lecture 12)

- Be able to distinguish each type of study design by its core defining features
- List the main strengths and weaknesses of each type of design
- Evaluate the appropriateness of each design for particular research questions
- Be able to interpret the findings from ecological studies, cross sectional surveys, case-control studies, cohort studies, meta-analysis, and randomised controlled trials.
- To understand the major sources of routine data on health and illness in the UK
- To be able to describe the strengths and weaknesses of routine health data
- to understand standardised mortality ratios and provide examples of their use in comparing health in populations

D. Sessions 6, 7 (lecture 13) and 8

- To be able to understand and interpret the statistical findings commonly reported in scientific papers
- To learn how to search for evidence from published medical journals on a specified research topic
- To learn how to read a paper in a scientific journal
- To be able to review and critically appraise the evidence presented in a scientific paper
- To be able to present critical appraisal findings to lecturers and peers

E. Sessions 5 (lecture 11) and 10 (lecture 17) - Disease Prevention

- To describe and give examples of the main methods of intervention to improve health (e.g. health education, health protection, and prevention)
- To describe and give examples of the different levels of disease prevention
- To understand the principles and practice of screening
- To be able to define validity for screening tests and calculate specificity, sensitivity and predictive value
- To understand the criteria for screening programmes

F. Session 9 - International Health

- To describe the extent of health and income inequalities worldwide
- To understand some of the key factors that might explain why some countries with similar incomes achieve variant child health outcomes
- To evaluate the reasons and solutions for health worker migration from poorer to richer countries
- To demonstrate an understanding of global health issues with regards to waterborne infectious disease

G. Session 10 (lecture 18)

- To revise key concepts in epidemiology and public national/international health

CONTACT DETAILS

Lecturers

Prof Sir Roy Anderson	Chair in Infectious Disease Epidemiology	roy.anderson@imperial.ac.uk
Dr. Sarah Fidler	Clinical Senior Lecturer in Communicable Diseases	s.fidler@imperial.ac.uk
Dr Mireille Toledano	Senior Lecturer in Epidemiology	m.toledano@imperial.ac.uk
Professor Paul Elliott	Head of the Department of Epidemiology and Biostatistics	p.elliott@imperial.ac.uk
Prof Majid Ezzati	Chair in Global Environmental Health	majid.ezzati@imperial.ac.uk
Dr Paul Aylin	Clinical Reader in Epidemiology & Public Health	p.aylin@imperial.ac.uk
Dr R Alex Bottle	Senior Lecturer in Medical Statistics	robert.bottle@imperial.ac.uk
Dr Petra Wark	Research Fellow	p.wark@imperial.ac.uk
Professor Helen Ward	Professor of Public Health	h.ward@imperial.ac.uk
Dr Tania Misra	Consultant in Communicable Disease Control, HPA	Tania.Misra@HPA.org.uk
Dr Teresa Norat	Principal Research Fellow	t.norat@imperial.ac.uk
Dr. Susan Hodgson	Lecturer Environmental Epidemiology & Exposures	susan.hodgson@imperial.ac.uk
Dr Claire Robertson	Senior Lecturer, Nutritional Epidemiology, Westminster University	c.robertson@westminster.ac.uk
Mr Mike Rowson	Senior Teaching Fellow, UCL Centre for International Health and Development	m.rowson@ucl.ac.uk
Prof Alan Fenwick OBE	Professor of Tropical Parasitology	a.fenwick@imperial.ac.uk
Dr. Bhargavi Rao	Welcome Clinical Research Fellow	bhargavi.rao@imperial.ac.uk

If you have any queries about the course, we have a dedicated email address for you to contact us on at eip@imperial.ac.uk. You can also email the course leader or any of the lecturers on the course directly (see above for details). Alternatively, you can leave a message at the Faculty Education Office (020 7594 9803; or email jo.williams@imperial.ac.uk).

Session 1: Lectures 1 and 2

AIDS: History and Progression
Global Health: Infectious Disease

Dr Sarah Fidler & Professor Sir Roy Anderson
(s.fidler@imperial.ac.uk; roy.anderson@imperial.ac.uk)

Learning Objectives

- Describe the current burden of infectious diseases and their disparities worldwide
- Identify the six commonest infectious causes of world mortality and some of the causes underlying their high incidence
- Define and distinguish incidence, prevalence, and mortality
- Understand the drivers of an AIDS epidemic, success and challenges of the response

---

Overview of session

Dr Sarah Fidler will review the achievements and challenges of the response to AIDS in developing countries, and discuss the issues for a long term response.

Prof Sir Roy Anderson will give a lecture on global health, focusing on infectious diseases. The lecture will discuss the changing global pattern of morbidity and mortality induced by infectious diseases over the past few decades. The relative importance of different diseases as causes of morbidity and mortality today will be detailed in both a global and local context. Changes in global patterns of infection will be discussed in terms of the rise in "global mixing" via various routes of transport including air traffic, the growth of 'mega cities' and the evolution of new infectious agents. Specific examples will be discussed, including tuberculosis, malaria, HIV/AIDS and dengue fever. Attentions will then turn to local conditions in the United Kingdom and chart the common infections and changing patterns therein. All these topics will be discussed within an epidemiological and evolutionary framework, and with reference to options for control at the level of the individual patient and the community (mass vaccination, anti-microbial and anti-viral treatment).

Notes

Burden of infectious disease: More than 90% of deaths from infectious diseases are caused by a handful of diseases: lower respiratory infections, HIV/AIDS, diarrhoeal diseases, tuberculosis, malaria and measles. Most notably, infectious diseases are the leading cause of death in sub-Saharan Africa (see chart).

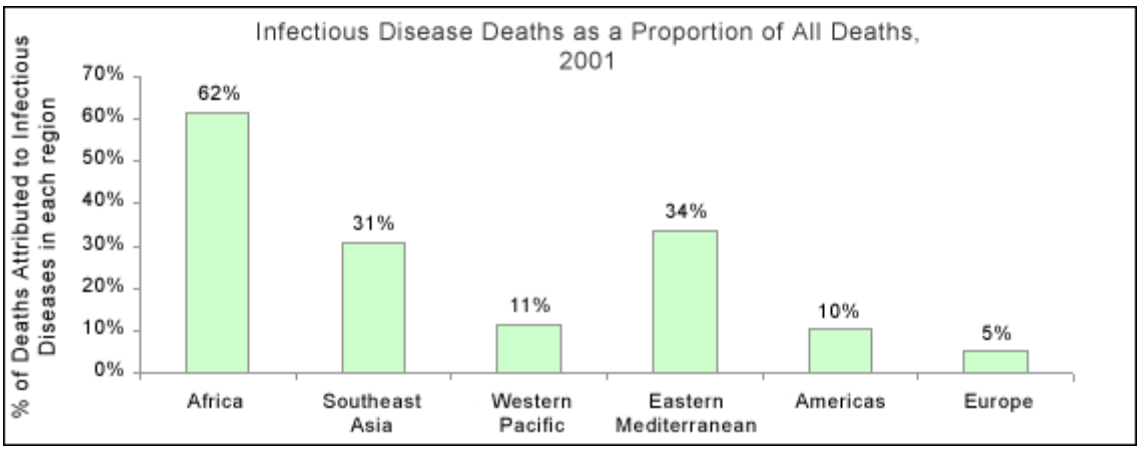


Figure 1. 10 leading causes of death, by income group

Figure 1 shows the ten most common causes of death globally by income group

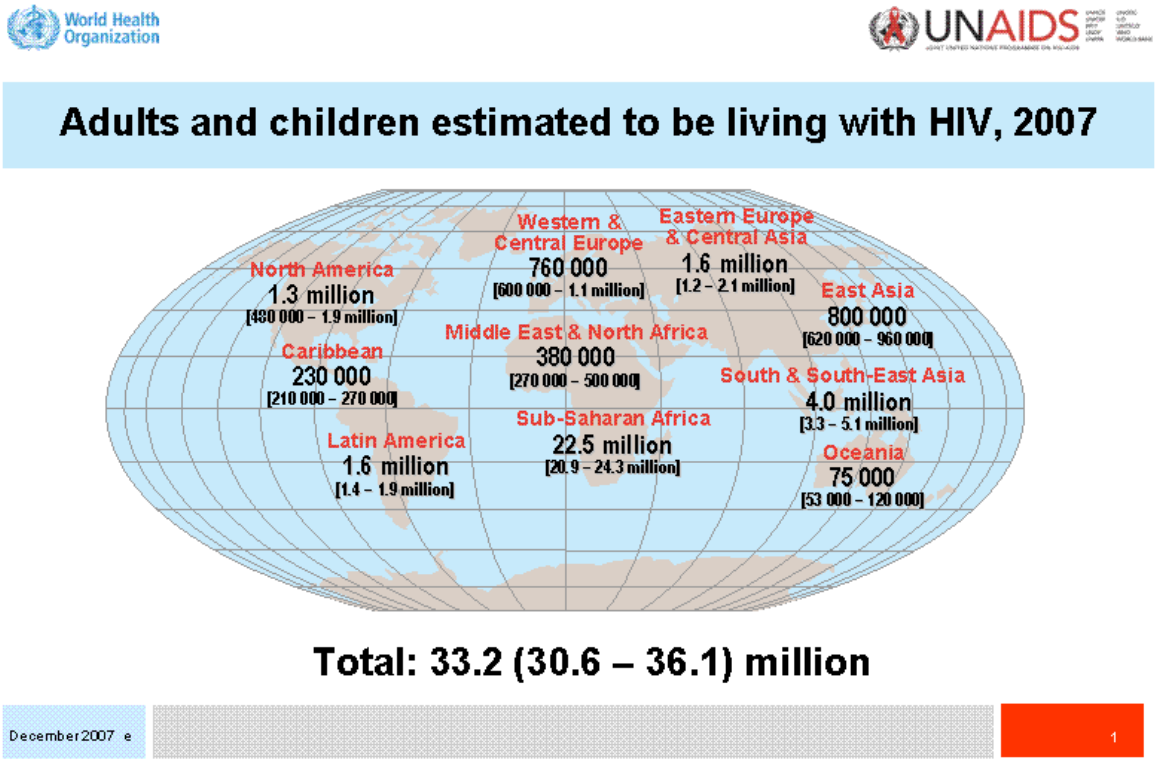
4 of the 10 leading causes of death are infectious diseases in low- and middle-income countries, 9 of the 10 leading causes are noncommunicable diseases in high-income countries

Leading Causes of Death Due to Infectious Diseases, 2002

Lower respiratory infections	3.9 million
HIV/AIDS	2.8 million
Diarrhoeal diseases	1.8 million
Tuberculosis	1.6 million
Malaria	1.2 million
Measles	0.6 million

Source: World Health Report, 2004 WHO

"...One place for disease is among the poor, esp when the poor are soci medically segregated those whose deaths be considered more im



http://data.unaids.org/pub/factsheet/2009/20091124_fs_global_en.pdf

Figure 2. The global distribution of adults and children living with HIV and AIDS in 2007

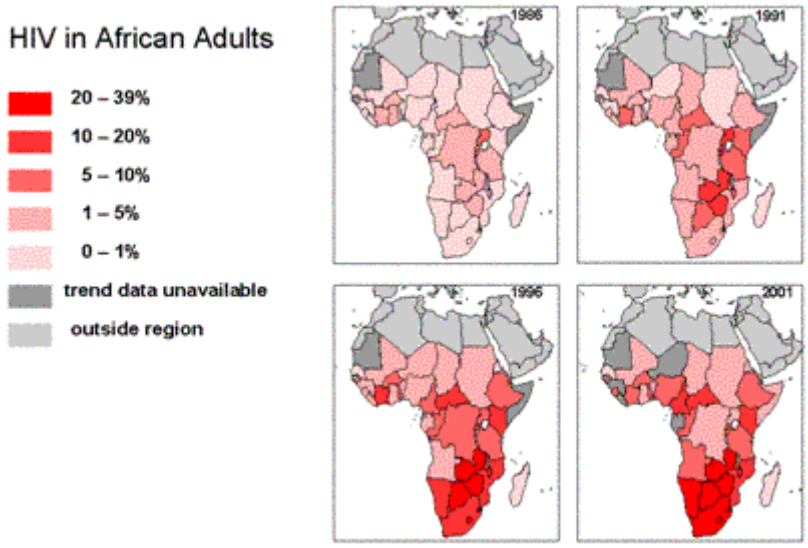


Figure 3 (above) shows the breakdown of the HIV prevalence in Sub-Saharan Africa between 1985-2001. Note that the epidemic is relatively recent especially in S Africa where in 1985 figures suggest the prevalence was < 5%

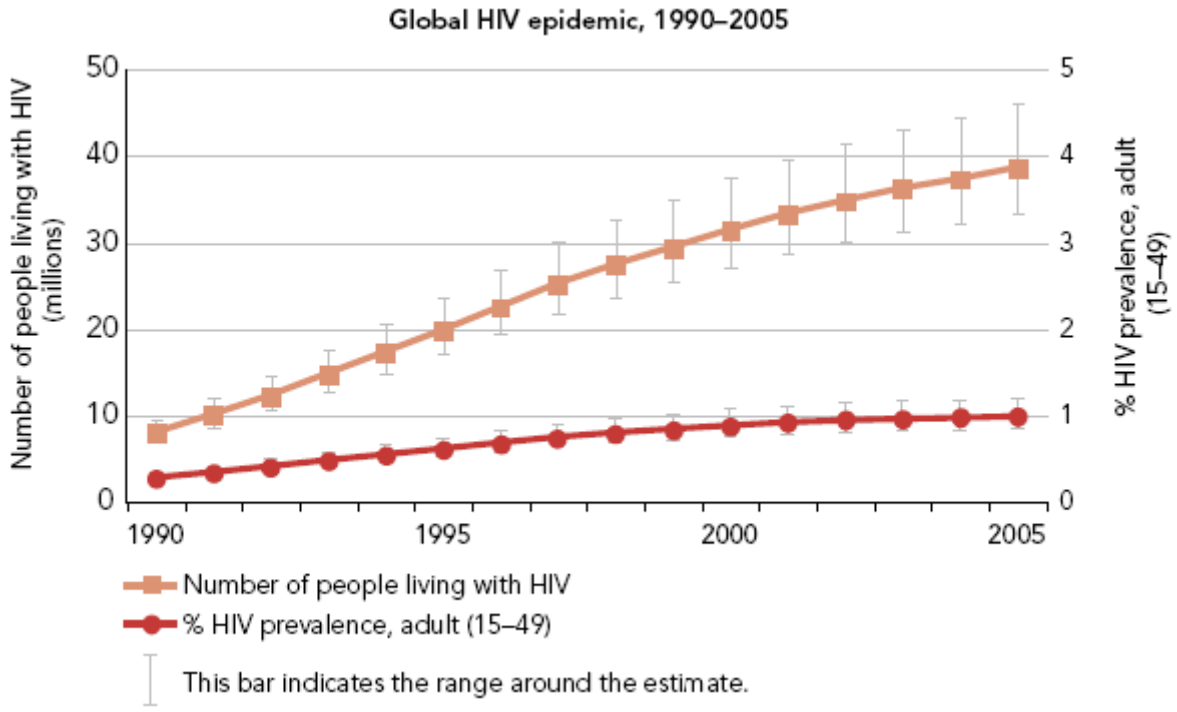
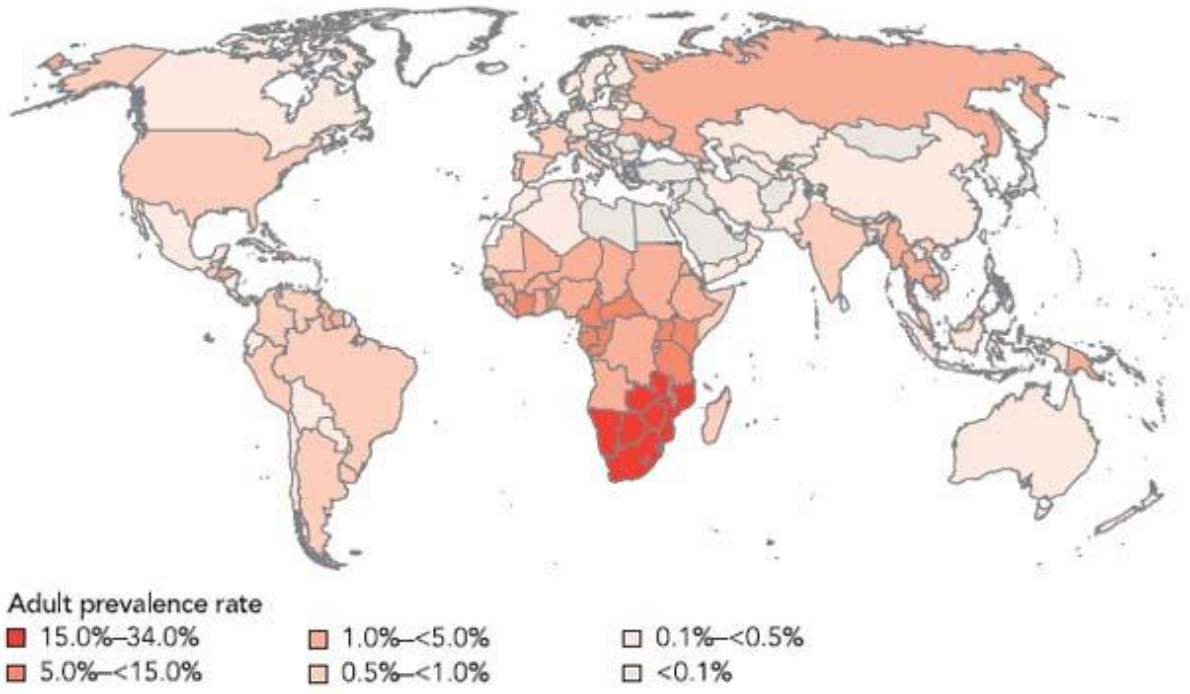


Figure 4 (above) shows the latest UNAIDS figures for the global HIV epidemic at last census. With the advent of ART the number of people living with HIV continues to increase leading to an increased prevalence



Similarly, Figure 5 (above) shows global estimates of adult prevalence in 2005

Age standardized death rates: Measuring how many people die each year and why they have died is one of the most important means – along with gauging how various diseases and injuries are affecting the living – of assessing the effectiveness of a country's health system. Having those numbers helps health authorities determine whether they are focussing on the right kinds of public health actions that will reduce the number of preventable deaths and disease. Globally, around 57 million people die each year. Almost 15% of these deaths occur in children under the age of 5. Most of these preventable deaths in children occur in low- and middle-income countries.

Under-5 mortality

7.6 million children under age five died in 2010, representing an under-five mortality rate of 57/1000 live births

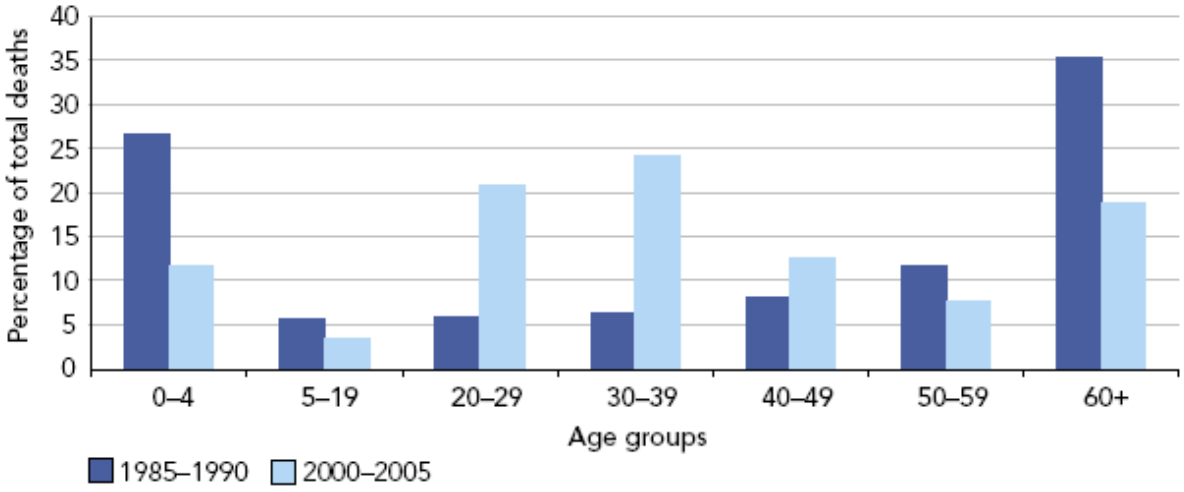


Figure 6 (above) is an outline showing the percentage of total deaths in SSA by age

comparing 2000 vs 2005 when ART became available

AIDS epidemic - successes and challenges of the response: One of the great success stories in the fight against AIDS, is the very broad access to antiretroviral therapy (ART) for HIV that has been achieved in poor countries. Here we see an illustration of the 10-fold growth of access to therapy for HIV/AIDS in Africa over a five-year period, from 2002 to 2007. HIV treatment is prolonging millions of lives but, unfortunately, we cannot treat our way out of this epidemic. For every person put on HIV treatment today, five are newly infected with HIV.

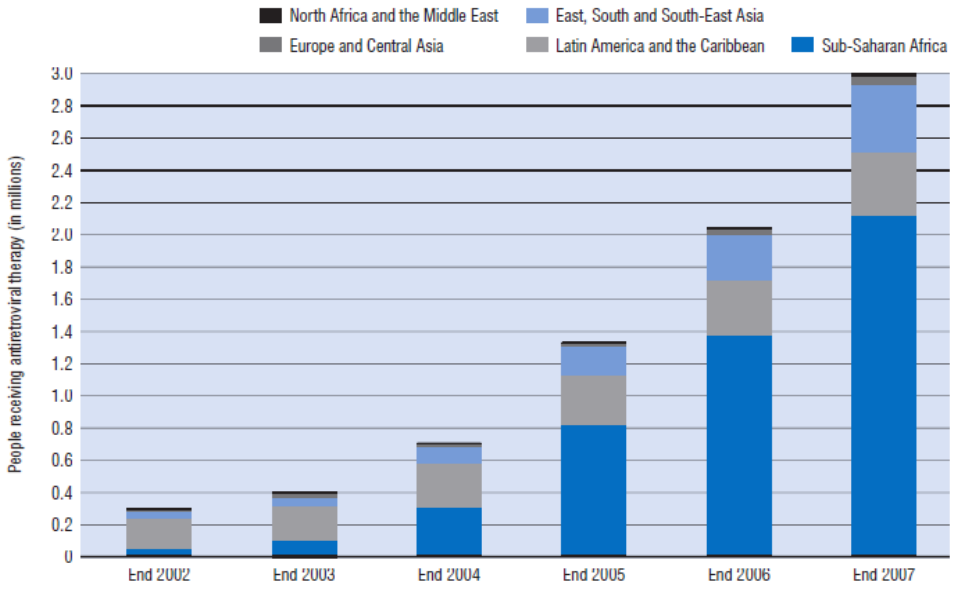
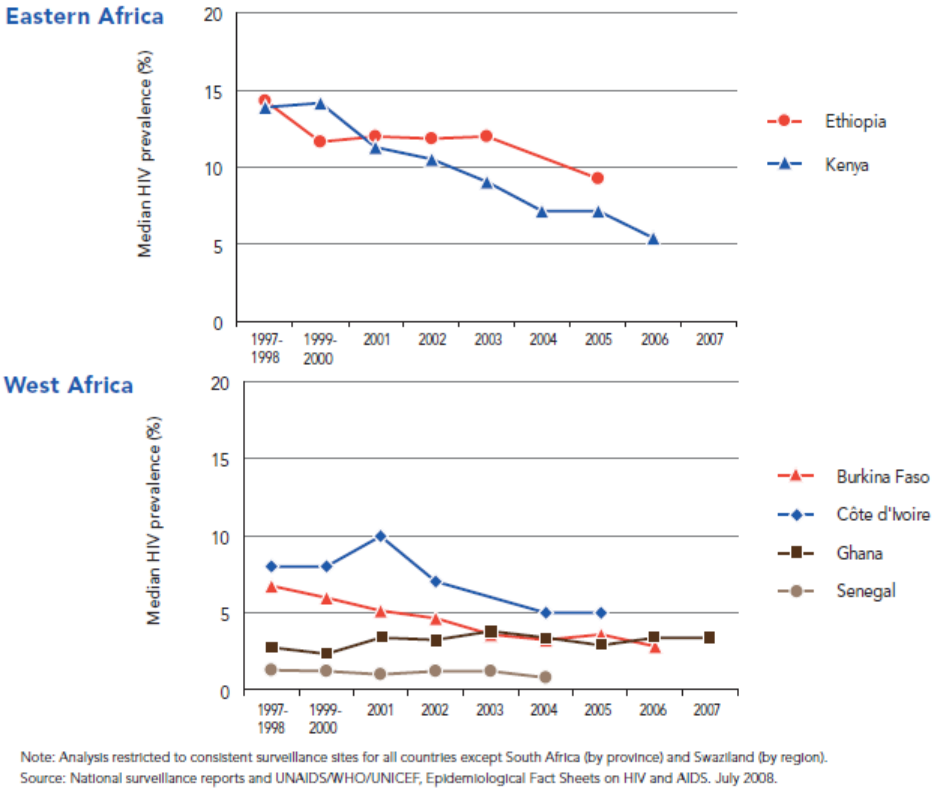


Figure 7 (above) shows the number of people globally receiving ART broken down by year

There have also been declines in HIV prevalence in pregnant women in recent years. There are a number of effective HIV prevention methods available today, including safer sex, safer injection practices, condom use, and male circumcision. There are, however, also social obstacles attached to each of these. And we have seen that, even when these interventions are fully funded and supported by states and social institutions, they have only been able to drive HIV infection rates down to a certain level.



Note: Analysis restricted to consistent surveillance sites for all countries except South Africa (by province) and Swaziland (by region).
Source: National surveillance reports and UNAIDS/WHO/UNICEF, Epidemiological Fact Sheets on HIV and AIDS. July 2008.

Figures 8 (above) show the reduction in HIV prevalence amongst women over time

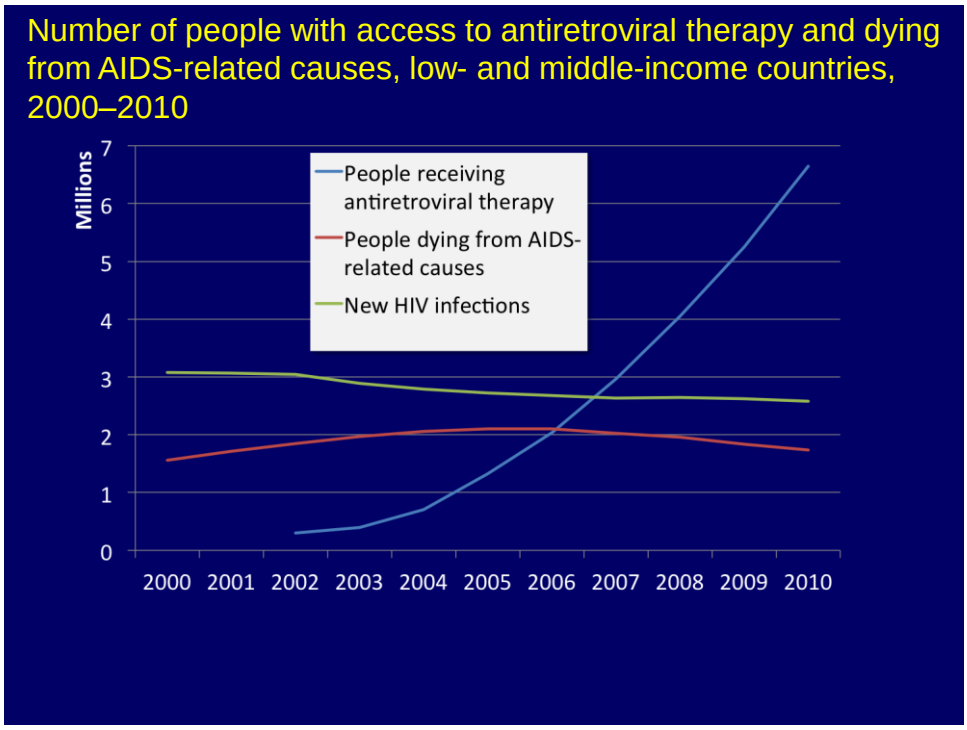


Figure 9 (above) shows WHO figures for the number of people with access to ART in low and middle income countries

In order to further reduce HIV incidence we need new biomedical tools -- the most important of these will be an effective HIV vaccine.

Key measures of disease in the population

Case

Epidemiology is based on the ability to quantify the occurrence of disease in populations. This requires a clear definition of what is meant by a case. This could be a person who has the disease, health disorder, or suffers the event of interest. The epidemiological definition of a case is not necessarily the same as the clinical definition**.**

Prevalence

Prevalence is the frequency of a disease in a population at a point in time; hence it is often called *point prevalence*

$$\text{Point prevalence} = \frac{\text{Number of cases in a defined population at one point in time}}{\text{Number of persons in a defined population at the same point in time}}$$

Prevalence is a proportion. It is the only measure of disease occurrence that can be obtained from cross sectional studies. It measures the burden of disease in a population. Prevalence measures status (a condition: a subject affected by a specific disease).

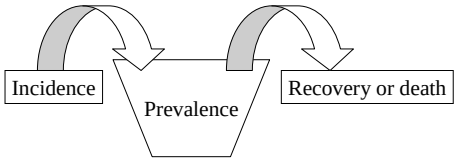
Incidence

Incidence quantifies the number of new cases of a disease within a specified time interval. Incidence measures events (a change from a healthy state to a diseased state).

$$\text{Incidence} = \frac{\text{Number of new cases of disease in a given time period}}{\text{Number of disease-free persons at the beginning of that time period}}$$

This measure of incidence can be interpreted as the probability, or risk, that an individual will develop the disease during a specific time period

Relationship of prevalence and incidence



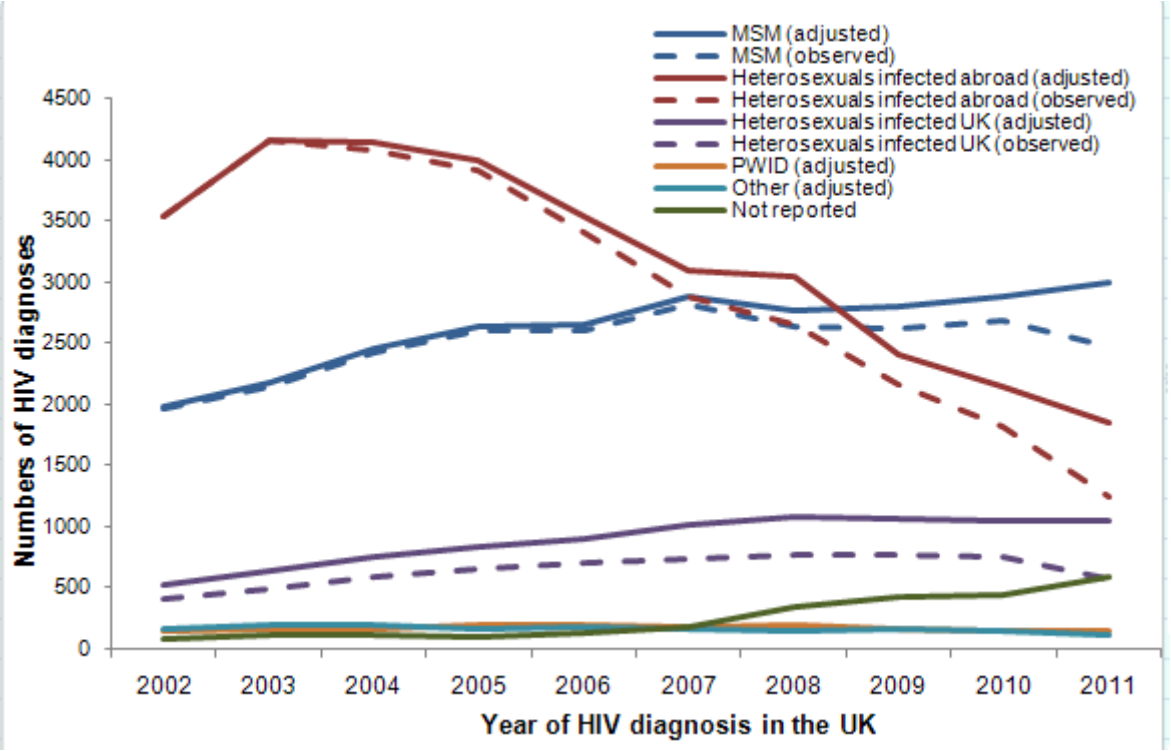
Prevalence = incidence x duration

Incidence measures new cases while **prevalence** measures all, cases new and old. The prevalence is dependent upon the number of new cases (incidence), and the time that they remain cases (duration of disease). Individuals only leave the “pool” of prevalent cases when they recover or die.

Example: HIV infection in the UK

In the UK, the numbers of new cases of HIV being diagnosed each year (incidence) is rising. The numbers of deaths from AIDS has declined, due to improved treatment with Highly Active Anti-Retroviral Therapy (HAART). Therefore the duration of disease is increasing. The consequence is a steep increase in the prevalence of HIV (the number of people living with HIV).

Figure 10. New HIV diagnoses by exposure group: United Kingdom, 2002 – 2011



Further reading/links

Vetter N, Mathews I, Epidemiology and Public Health Medicine, pp23-29.

Coggan et al. Epidemiology for the Uninitiated, http://www.bmj.com/epidem/epid.html. Chapter 2.

McNeill, W. (1976). Plagues and People. Blockwell, Oxford.

Anderson, R.M. and May, R.M. (1991). Infectious Diseases of Humans: dynamics and control. Oxford University Press, Oxford.

Shilts R. And the Band Played On: politics, people and the AIDS epidemic. Penguin Books, 1987

Piot P, Bartos M, Larson H, Zewdie,D, Mane P. Coming to terms with complexity: a call to action for HIV prevention. Lancet, 2008; 372:845-59

Piot P, Russell S, Larson H. Good Politics, Bad Politics: The Experience of AIDS. American Journal of Public Health 2007; 97: 1934-36

Piot P. AIDS: from crisis management to sustained strategic response. Lancet 2006;368:526-30

Please note that this information may be subject to change

Session 2: Lectures 3 and 4

Global Health: Non-infectious Disease

Professor Paul Elliott and Professor Majid Ezzati
(p.elliott@imperial@imperial.ac.uk; majid.ezzatti@imperial.ac.uk)

Learning Objectives

- Explain the concept of epidemiological transition
- Describe the current burden of non-communicable diseases and their disparities worldwide.
- Identify the commonest non-infectious causes of world mortality and some of the causes underlying their high incidence

---

Overview of session

The first half of this lecture will focus on the global epidemiology of cardiovascular disease and the second half will focus on global epidemiology of cancer.

Global health: Epidemiological transition

Epidemiological transition – this is the changes in levels and causes of mortality, which is commonly summarized as a decline in total mortality, and a significant reduction in infectious and deficiency diseases, which increase the relative role of chronic non-communicable diseases like cancers, cardiovascular and chronic respiratory disease, and diabetes. It accompanies socio-demographic and health system changes among the poorer countries but continues in more industrialized nations. With advances in clinical medicine and epidemiology, it has become apparent that this transition is complex and dynamic: the health and disease patterns of a society evolve in diverse ways as a result of demographic, socioeconomic, technological, cultural, environmental and biological changes. It is rather a continuous transformation process, with some diseases disappearing and others appearing or re-emerging. There are some outstanding examples, such as the emergence of new infectious diseases like AIDS, the increase in infections that were previously controlled, such as tuberculosis and dengue fever, the decline in stomach cancer and the rise and fall of lung cancer, and the shift from stroke to heart disease.

Global health: cardiovascular disease epidemiology

Cardiovascular diseases (mainly coronary heart disease and stroke) accounted for some 14.3 million deaths worldwide (in 1990), 28.3% of all deaths. Many more such deaths occurred in the developing world (9.1 million) than the developed world (5.2 million). Coronary heart disease and stroke respectively rank first and second among cause-specific mortality worldwide. Because of the demographic and epidemiological transitions, the burden of disease from non-communicable diseases in the developing countries is likely to rise: an estimated more than doubling of mortality from both coronary heart disease and stroke in developing countries, comprising an estimated 69% and 76% respectively of all deaths from these causes worldwide.

There are wide discrepancies in incidence and mortality from coronary heart disease, having low rates in Japan and high rates in the UK and other western countries, and in the formerly socialist economies of Europe. At all ages, rates are higher in men than women. Trends in both coronary heart disease and stroke mortality have been declining in many countries in recent years (after a large rise in coronary heart disease mortality up to the 1960s and 1970s), though there has been a recent rise in the formerly socialist economies of Europe. These epidemiological patterns (rising and declining rates within countries, large differences across countries which lessen or disappear with migration) indicate that environmental rather than genetic factors underlie much of the variation in cardiovascular disease risk worldwide.

Three risk factors related to diet and lifestyle (and therefore modifiable) are particularly important: high blood pressure, tobacco smoking and serum cholesterol levels. Again the

burden of disease attributable to these risk factors is high in the developing as well as the developed countries. Worldwide trends in overweight and obesity will increase the burden of Non-communicable disease including metabolic disorders and diabetes.

Global health: cancer epidemiology
Cancer is a major public health problem throughout the world, causing more than a quarter of all deaths in many countries. Cancer accounted for about 12.5% of the deaths worldwide in 2002, when 11 million people were diagnosed with cancer. By 2020, there could be as many as 15 million new cases per year. Cancer burden is shifting to less developed countries, in which 60% of these cases are likely to occur.

Many types of cancer vary in incidence between different populations and every type of cancer is rare in some parts of the world. Lung, breast and colorectal cancer are currently the most commonly diagnosed cancers, whereas lung cancer, stomach cancer and liver cancer are the most common causes of cancer death. Cancer rates in migrants tend to converge towards local cancer rates over time, pointing to a role for modifiable risk factors. At least a third of all cancers are likely to be preventable.

Age-specific cancer incidence and mortality rates have fallen for some cancer sites, while other cancers have become more common, reflecting changes in relevant exposures, diagnosis, treatment, and screening. Because cancer can take 20 years to appear, current cancer rates are affected by changes and exposures that took place in the past. Rates of smoking-related cancers in women, for example, will continue to increase in most countries; as will the number of cases attributable to asbestos exposure. Smoking and overweight may become more important contributors to cancer rates than infections in some countries.

The lecture will cover global patterns of cancer, to a great extent building on the International Agency for Research on Cancer (IARC) Globocan information (available via the Internet at http://www-dep.iarc.fr/). The importance of prevention will also be stressed, using examples from various parts of the world. Some of the major known carcinogens will be discussed (tobacco, alcohol, air pollution and occupational agents), but also infections, diet and obesity.

Further reading

Epidemiological transition

Salomon JA, Murray CJL. The epidemiologic transition re-examined: compositional models for causes of death by age and sex. *Population and Development Review 2002;* 28:205-228

Lopez AD, Mathers CD, Ezzati M, Jamison D, Murray CJL. The global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 2006; 367(9524):1747-1757

Ezzati M, Lopez AD, Rodgers A, Vander Hoorn S, Murray CJL and the Comparative Risk Assessment Collaborating Group. Selected major risk factors and global and regional burden of disease. *Lancet* 2002; 360: 1347-60.

Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Medicine* 2006; 3(11):e442

Cardiovascular epidemiology:

Marmot M, Elliott P, eds. *Coronary Heart Disease Epidemiology: From Aetiology to Public Health.* Oxford: Oxford University Press 2005, 932 pp.

Ezzati M, Vander Hoorn S, Lawes CMM, Leach R, James WPT, Lopez AD, Rodgers A, Murray CJL. Rethinking “the diseases of affluence” paradigm: global patterns of nutritional risks in

relation to economic development. *PLoS Medicine* 2005; 2(5):e133

Kuulasmaa K, Tunstall-Pedoe H, Dobson A, Fortmann S, Sans S, Tolonen H, Evans A, Ferrario M, Tuomilehto J. Estimation of contribution of changes in classic risk factors to trends in coronary-event rates across the WHO MONICA Project populations. *Lancet* 2000; 355:675-687

Danaei G, Finucane MM, Lin JK, Singh GM, Paciorek CJ, Cowan MJ, Farzadfar F, Stevens GA, Lim SS, Riley LM, Ezzati M on behalf of the Global Burden of Metabolic Risk Factor of Chronic Diseases Collaborating Group (Blood Pressure). National, regional, and global trends in systolic blood pressure since 1980: Systematic analysis of health examination surveys and epidemiological studies with 786 country-years and 5.4 million participants. *Lancet* 2011; 377(9765):568-577

Cancer epidemiology:

Danaei G, Vander Hoorn S, Lopez AD, Murray CJL, Ezzati M, the Comparative Risk Assessment Collaborating Group. Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors. *Lancet* 2005; 366(9499):1784-1793

Parkin DM, Bray F, Ferlay J, and Pisani P (2005) *Global cancer statistics, 2002.* CA: A Cancer Journal for Clinicians 55: 74–108.

Parkin DM (2006) The global health burden of infection-associated cancers in the year 2002. *Int J Cancer* 118: 3030–44.

Lin HH, Murray MM, Cohen T, Colijn C, Ezzati M. Effects of smoking and solid-fuel use on COPD, lung cancer, and tuberculosis in China: a time-based, multiple-risk-factor modelling study. *Lancet* 2008; 372(9648):1473-1483

Wark PA, Peto J. Cancer Epidemiology. In: Heggenhougen K, Quah S, eds. *International Encyclopedia of Public Health, Vol 1.* San Diego: Academic Press; 2008. pp 416-24.

Session 3: Lectures 5, 6 and 7

The importance of evidence in the practice of medicine

Dr Paul Aylin
(p.aylin@imperial.ac.uk)

Learning Objectives

- Recognise the role of evidence based practice in clinical medicine
- List and define possible explanations for observed associations (chance, bias, confounding, causation), and cite examples of each
- Be able to describe the hierarchy of evidence in study design
- List the Bradford-Hill criteria for establishing causation and apply these to specific examples
- Be able to apply epidemiological skills to clinical decision making

---

Evidence-based medicine

- The concept of evidence based medicine has been evolving over the past 30 years.
- Methods to critically appraise clinical information and classify it according to the strength of evidence was first presented in a Canadian Medical Association Journal series on how to critically appraise literature in the early 1980s.
- Concepts emerging from the literature on “critical appraisal” promoted what has become known as evidence based medicine (EBM), suggesting that clinicians should use critically appraised information in clinical practice for optimal care of their patients

Criticism of Evidence based medicine

- It is impossible for any clinician to have the time to critically appraise even one article per week let alone the number that would need to be appraised to answer questions (estimated at 3.5 per clinical session) arising in a busy practice.
- Governments, healthcare commissioners and providers have used the jargon of EBM to justify decisions, directives, or incentives that are seen by clinicians as inappropriate

Why EBM matters to Clinicians

- Revalidation
- Patient Care
- Medical Knowledge
- Practice-Based Learning and Improvement
- Interpersonal and Communication Skills
- Professionalism

Evidence based medicine does NOT replace clinical decision making but is only a tool

Hierarchy of studies

- Systematic reviews and meta-analyses
- Randomised Controlled Trials
- Cohort studies
- Case-control studies

- Ecological studies
- Descriptive/cross-sectional studies
- Case report/series

Association and causation
Association refers to the statistical dependence between two variables, that is the degree to which the rate of disease in persons with a specific exposure is either higher or lower than the rate of disease without that exposure.

A link, relationship or correlation

Evaluating a statistical association

Consider chance, bias, confounding, cause

Chance
Make inference from samples rather than whole populations
- Sample size
- Power calculations
- P values and statistical significance

Bias
A systematic error
- Selection bias
- Measurement bias
- Observer bias
- Responder bias

Confounding
Mixing of effects between exposure, the disease and a third factor
Account for confounding using matching, randomisation, stratification and multivariate analysis

Causal effect
Judgement of a cause-effect relationship
Judgement based on a chain of logic that addresses two main areas:
- Observed association between an exposure and a disease is valid
- Totality of evidence taken from a number of sources supports a judgement of causality

Factors to consider:

1. Strength

The strength of an association is measured by the magnitude of the relative risk. A strong association is more likely to be causal than is a weak association, which could more easily be the result of confounding or bias. However, a weak association does nor rule out a causal connection. For example, passive smoking and lung cancer.

2. Consistency

If similar results have been found in different populations using different study designs then the association is more likely to be causal since it is unlikely that all studies were subject to the same type of errors. However, a lack of consistency does not exclude a causal association since different exposure levels and other conditions may reduce the impact of the causal factor in certain studies.

3. Specificity

If a particular exposure increases the risk of a certain disease but not the risk of other diseases then this is strong evidence in favour of a cause-effect relationship e.g. Mesothelioma. However, one-to-one relationships between exposure and disease are rare and lack of specificity should not be used to refute a causal relationship; for example cigarette smoking causes many diseases.

4. Temporal relationship

This is an essential criterion. For a putative risk factor to be the cause of a disease it has to precede the disease. This is generally easier to establish from cohort studies but rather difficult to establish from cross-sectional or case-control studies when measurements of the possible cause and the effect are made at the same time. However, it does not follow that a reverse time order is evidence against the hypothesis.

5. Dose-response relationship

Further evidence of a causal relationship is provided if increasing levels of exposure lead to increasing risks of disease. Some causal associations, however, show a single jump (threshold) rather than a monotonic trend.

6. Plausibility

The association is more likely to be causal if consistent with other knowledge (e.g. animal experiments, biological mechanisms, etc.). However, this criterion should not be taken too seriously because lack of plausibility may simply reflect lack of scientific knowledge. The idea of microscopic animals or animalcules as cause of disease was distinctly implausible before Van Leeuwenhoek’s microscope

7. Coherence

Coherence implies that a cause and effect interpretation does not conflict with what is known of the natural history. However absence of coherent information as distinguished from the presence of conflicting information, should not be taken as evidence against an association being causal.

8. Experimental evidence

Experimental evidence on humans or animals. Evidence from human experiments is seldom available and animal research relates to different species and different levels of exposure.

9. Analogy

At best analogy provides a source of more elaborate hypotheses about the association in question. Absence of such analogies only reflects lack of imagination or experience, not falsity of the hypothesis (Bradford Hill 1965).

Further reading

Ben Goldacre (2009). Bad Science. Harper Perennial ISBN-13: 978-0007284870

Trisha Greenhalgh (2006 3rd Edition edition). How to Read a Paper: The Basics of Evidence Based Medicine. WileyBlackwell;

Clinical evidence
Available at url: http://www.clinicalevidence.com/ceweb/conditions/index.jsp

Austin Bradford Hill, “The Environment and Disease: Association or Causation?,”Proceedings of the Royal Society of Medicine, 58 (1965), 295-300.
Available at url: http://www.edwardtufte.com/tufte/hill

Session 4: Lecture 8

Study Design: Descriptive studies and routine data

Dr Alex Bottle
(robert.bottle@imperial.ac.uk)

Learning Objectives

- To understand the major sources of routine data on health and illness in the UK
- To be able to describe the strengths and weaknesses of routine health data
- To understand the construction of standardised mortality ratios (SMRs) and how they can be used in comparing health in populations

---

Hierarchy of study design

- Systematic reviews and meta-analyses (highest – but can still be inadequate)
- Randomised Controlled Trials
- Cohort studies
- Case-control studies
- Ecological studies
- Descriptive/cross-sectional studies
- Case report/series (lowest – but can still be valuable)

Descriptive studies in epidemiology examine the distribution of disease across various factors including population or sub-groups, geographical location and time period

Cross sectional survey examples

- 2001 Census
- Health Survey for England
- NHS Inpatient Survey on patient experience

Routine data

- "Data that are routinely collected and recorded in an ongoing systematic way, often for administrative or statutory purposes and without any specific research question in mind at the time of collection" (Hansell A, Aylin P. Using routine data in health impact assessment. JPHM 2001)

Types of routine data

- Health outcome data, e.g. deaths, hospital admissions and primary care consultations or prescriptions, levels of well-being from national surveys
- Exposures and health determinant data, e.g. smoking, air pollution, crime statistics
- Disease prevention data, e.g. screening and immunisation uptake
- Demographic data, e.g. census population counts
- Geographical data, e.g. health authority boundaries, location of GP practices
- Births
- Deaths
- Cancer registrations
- Notifications of infectious diseases
- Terminations of pregnancy
- Congenital anomalies

- Hospital admissions
- Community systems
- GP consultation data
- Prescriptions
- Road Traffic Accidents

Standardised Mortality Ratio

Key confounders such as age may vary between populations. Population death rates may be compared taking into account (or "adjusting") for the effect of age. One method for comparing rates is the Standardised Mortality Ratio (SMR). The SMR is a rate ratio adjusted for age. It represents the ratio of the number of observed deaths (or cases of disease) (O) in a particular population to the number that would be expected (E), if that population had the same mortality or morbidity experience as a standard population, corrected for differences in age structure.

$$\text{SMR} = \frac{\text{Number of observed deaths}}{\text{Number of expected deaths if experienced the same age specific rates as standard population}}$$

It is common for SMRs to be adjusted for age and also for sex.

Further reading

Trisha Greenhalgh (2001). How to Read a Paper: The Basics of Evidence Based Medicine. BMJ Books

Session 4: Lecture 9

Study Design: Cohort and Case Control Studies

Dr Petra Wark
(p.wark@imperial.ac.uk)

Learning Objectives

- To distinguish and describe the design of case control and cohort studies by their core defining features
- To describe where cohort and case control studies fit in the hierarchy of epidemiological studies
- To list the strengths and weaknesses of cohort studies and case control studies
- To be able to interpret odds ratios and rate ratios
- To be able to calculate crude odds ratios and rate ratios from a two-by-two table
- To be able to evaluate the appropriateness of case control and cohort designs for particular research questions

---

Hierarchy of study design (high to low)

- Systematic reviews and meta-analyses
- Randomised Controlled Trials (experimental studies)
- ***Cohort studies***
- ***Case-control studies***
- Ecological studies
- Descriptive/cross-sectional studies
- Case report/series

Cohort Studies

- Observational analytical epidemiological studies
- A group of people (cohort) followed over time
- Prospective or retrospective design
- A prospective cohort study ascertains disease during follow-up, whereas a retrospective cohort study looks at events that already happened
- Exposures measured prior to disease (prospective design)
- Retrospective cohort studies use previously recorded information on exposure
- Can directly measure incidence of disease in exposed and non-exposed people, which information can be used to calculate rate ratios or risk ratios

Strengths of cohort studies

- Able to look at multiple outcomes
- Incidence (number of new cases in a defined time period) can be calculated
- Good to look at rare exposures
- Causal effect can be studied in prospective design

Weaknesses of cohort studies

- Time-consuming (prospective design)
- Expensive (prospective design)
- Loss to follow up may introduce bias
- Healthy worker effect may cause bias in occupational cohorts
- Inefficient for studying rare diseases

Case control studies

- Observational analytical epidemiological studies
- Retrospective design
- Cases are defined and their exposure compared with controls
- Controls (free of disease) are selected to represent source population of cases
- Exposure determined post-diagnosis
- The odds ratio is the only measure of relative risk that can be calculated

Strengths of case control studies

- Relatively quick and inexpensive
- Good at examining diseases with long latency periods
- Good design to evaluate rare diseases
- Can examine effects of multiple exposures

Weaknesses of case control studies

- Prone to bias – particularly selection bias and recall bias
- Inefficient to examine effects of rare exposures
- Cannot calculate incidence rates directly
- Temporal relationship between exposure and disease is hard to establish

Further Reading

- Hennekens CH, Buring JE. Epidemiology in Medicine. Little, Brown & Co. 1987.
- Porta M, Last JM. A Dictionary of Epidemiology. Oxford University Press. (5th Edition). 2008.
- Rothman KJ. Epidemiology: An introduction. Oxford University Press Inc 2002, USA
- Dos Santos Silva I. (Ed) Cancer Epidemiology : Principles and Methods. IARC Lyon France (Available in French, English, Spanish.
- Coggon D, Barker DJP, Rose G. (Ed). Epidemiology for the uninitiated. BMJ Books 4rd edition, 2003.

Session 5: Lecture 10

More on Study Design: Clinical Trials

Professor Helen Ward
(h.ward@imperial.ac.uk)

Learning Objectives

- To understand the unique significance of, and key components in, the clinical trial design
- To appreciate the potential biases and limitations in clinical trials
- To be able to interpret the findings presented from clinical trials
- To be able to evaluate the appropriateness of the clinical trial design for particular research questions

---

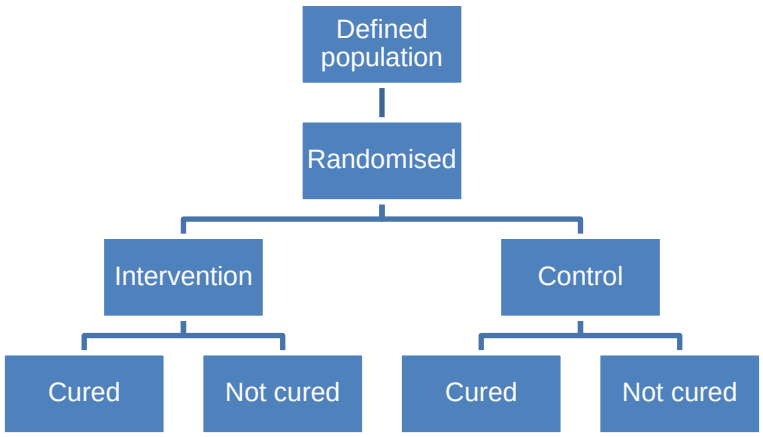
What is a clinical trial and why is it unique?

A clinical trial is a planned **experiment** in humans, designed to measure the effectiveness of an intervention. The intervention is usually a new drug, but the method can equally be applied to the assessment of a surgical procedure, a vaccine, complementary therapy etc.
Experimental studies like trials are different from most epidemiological studies (surveys, cross sectional, cohort, case control, ecological) which are **observational**. In observational studies the investigator measures what happens but does not control it. For example, an investigator may record whether people smoke, and relates this to whether or not they develop lung cancer. In contrast, in a clinical trial, the investigator would allocate one group to smoking and the others to not smoking, and then see who got ill. Of course in this example this would not be done as it is both unethical and impractical since people do not smoke or not just because someone tells them to.

Features of a clinical trial

- **Experimental** study
- Must contain a **control** group
- **Prospective**: participants are followed through time
- Patients are enrolled, treated and followed over same period of time
- Participants should be **randomised** to control or intervention groups
- Ideally the participants and the researcher are unaware if a participant has been assigned to the treatment or control group. This is known as **blinding**.

Design:



Why a control group?
The control group is those study participants who do not receive the intervention under assessment. A control group must be included otherwise you cannot be sure why the outcome happened; it may be due to the new treatment or it may have happened anyway. Control groups may be given a placebo (an inactive substance such as a sugar pill, or water injection), or a standard treatment.

Why randomise?
People who are eligible for the trial (i.e. have the condition you are interested in) are recruited, consent is obtained and then they are randomly allocated to the intervention or control groups. Randomisation is done to remove treatment **allocation bias**. Without randomisation it is possible (indeed likely) that the investigator will choose different patients for each group. In a famous early study of the BGC vaccine for TB in children, deaths from TB were five times higher in the control group than the vaccinated children.[1] Further investigation showed that doctors had tended to offer the new vaccine to children whose parents were more ”cooperative”, and left the rest as controls. These cooperative parents were likely to have been more educated, health conscious and therefore to have a lower mortality from TB regardless of the vaccination. Hence the need for randomisation in treatment allocation.

Why blind or double-blind?
Blinding means that the patient does not know whether they are getting the new treatment or not. In a double blind trial neither the patient nor the doctor knows which treatment they are getting. This is to prevent bias in reporting or measurement of the outcome, **measurement bias**. People who are getting a new treatment (or treatment compared with no treatment) often report improvement in subjective symptoms because they are enthusiastic and hopeful. Similarly if a doctor knows that a patient is on the new or active drug they may look for more improvements.

Ethics and consent
Clinical trials are strictly regulated to ensure that patients are protected. All clinical trials have to be registered, reviewed by an independent scientific committee, be approved by a Research Ethics Committee and adhere to government and international guidelines. Trials will have an independent data monitoring committee – a group of independent researchers who can check progress during the trial; they will usually un-blind the results to see if there is any major difference in outcome (improvement or side effects) between the intervention and control groups. If there is a large difference they have the power to stop the trial.
All participants in a trial must provide informed consent, and be free to withdraw at any time without affecting their care.

Analysis, evaluation and reporting

At the end of a trial the results will be analysed. How these are presented will depend on the particular design of the study. The outcomes are presented in terms of **efficacy** (the true biological effect of a treatment) or **effectiveness** (effect of a treatment when actually used in practice). Trial outputs:

The experimental event rate (EER) = incidence in the intervention arm

Control event rate (CER) = incidence in the control arm

Relative risk = EER/CER

Relative reduction = (CER- EER)/ CER

Absolute risk reduction (ARR) = CER- EER

Number needed to treat (NNT) = 1/ARR

Reporting trials

Clinical trials are now expected to be reported according to the CONSORT (Consolidated Standards of Reporting Trials) guidelines.This ensures that papers about trials include all the relevant information for readers to critically appraise the paper.

Phases of clinical trials

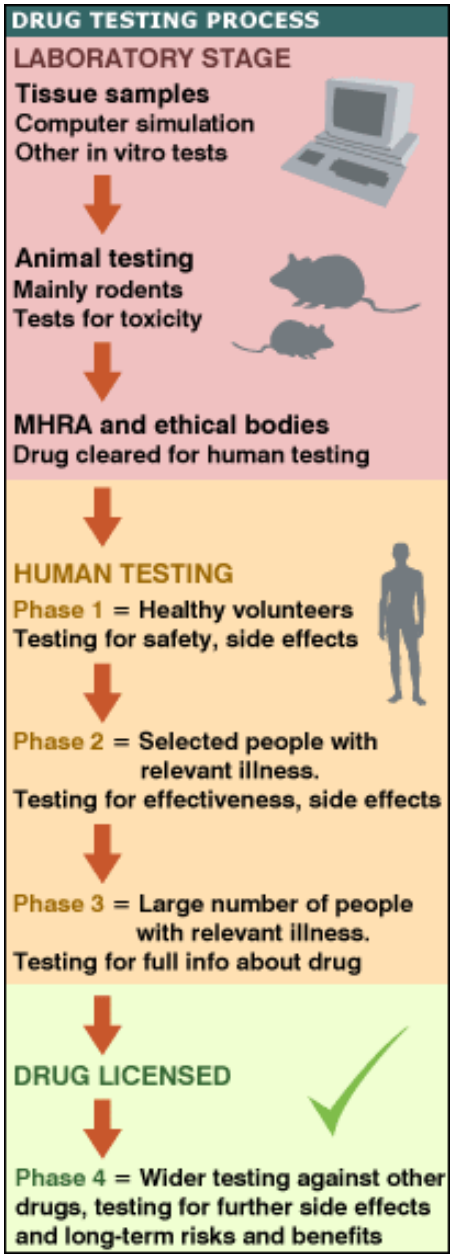
Several stages must be followed in the development and evaluation of a new drug to ensure it is safe and effective.[3]

Phase I trials aim to test the safety of a new treatment. This will include looking at side effects of a treatment – for example, does it make people sick, raise their blood pressure etc? Phase I trials involve only a small number of people, usually healthy volunteers. The disastrous trial of the anti-inflammatory drug TGN1412 at Northwick Park Hospital in 2006 was a phase I study.

Phase II trials test the new treatment in a larger group of people who have the disease for which the treatment is to be used, to see whether the treatment is promising, i.e. effective at least in the short term. Usually a few hundred people are involved at this stage. Phase II trials also look at safety.

Phase III trials test the new treatment in a larger group of people. Phase III trials compare the new treatment with the treatment currently in use, or with a placebo. These trials look at how well the new treatment works, and at any side effects it may cause. Often several thousand patients will be involved in a phase III trial. They may use different hospitals and live in different countries. The MRC Clinical Trials Unit runs phase III trials across Europe, parts of Africa and the USA. The smaller the expected advantage of a treatment, the more people will be needed to take part in a trial.

Phase IV trials are done after the drug or treatment has been marketed to gather information on the drug's effect in various populations and any side effects associated with long-term use.



Example: ACCOMPLISH hypertension treatment trial[1]

An RCT was carried out on patients with hypertension, comparing benazepril plus amlodipine (the intervention) to benazepril plus hydrochlorthiazide (the control). The primary endpoint was a vascular event (death from cardiovascular causes, non-fatal myocardial infarction, non-fatal stroke etc).

Treatment	Number	Events	Event rate
Intervention	5744	552	9.6%
Control	5762	679	11.8%

[1] Jamerson K et al. *N Engl J Med* 2008;359:2417-28

Experimental event rate (EER) = 9.6%

Control event rate (CER) = 11.8%

Relative risk reduction = (CER – EER)/CER = (11.8 – 9.6)/11.8 = 18.6%
(there was a 19% reduction in events in people taking the new treatment compared with controls)

Absolute risk reduction (ARR) = CER- EER = 2.2%
(The risk of an event was 2.2% lower in those taking the new treatment)

Number needed to treat (NNT) = 1/ARR = 45
(You would need to treat 45 people with the new treatment for an average of three years to avoid one additional vascular event)

Recommended reading and further examples

Ward H, Toledano M, Shaddick G, Davies B, Elliott P. *Oxford Handbook of Epidemiology for Clinicians*. The following sections deal with clinical trials and their interpretation:

p 42 How effective is the treatment
p 44 Interpreting reports of clinical trials
p 82 Appraisal checklist
p 84 CONSORT statement
p 186 Intervention studies and clinical trials
p 190 Clinical trials: examples
p 192 Clinical trial phases

Session 5: Lecture 11

Disease Prevention I: Public Health and Health Promotion

Dr Tania Misra
(Tania.Misra@HPA.org.uk)

Learning outcomes

- Familiarity with the core concepts of public health and health promotion
- Definitions
- Wider determinants of health
- High Risk vs. Population approach
- Conceptual framework for designing Public Health and Health Promotion interventions
- Knowledge of examples of evidence based programmes – Public Health interventions in the UK and abroad

Health

The word "health" is derived from an Old English word 'hael', which means 'whole'. WHO defines health as:

"A resource for everyday life, not the objective of living. Health is a positive concept emphasising social and personal resources, as well as physical capacities".

Public Health

Definition: The science and art of preventing disease, prolonging life and promoting health through organised efforts of society

Illustration of Indicators of health

Several slides on indicators of health
The wider determinants of health

Health Promotion (empowering people for health)

Health Promotion is the process of enabling people to increase control over, and to improve their health. (Ottawa Charter for Health Promotion, WHO. Geneva, 1986)

Health promotion represents a comprehensive social and political process, it not only embraces actions directed at strengthening the skills and capabilities of individuals, but also action towards changing the social, environmental and economic conditions so as to alleviate their impact on public and individual health. Health promotion is the process of enabling people to increase control over the *determinants of health* and thereby improve their *health.* Participation is essential to sustain health promotion action.

Health Promotion involves:

- Clinical intervention

  Biomedical (classically thought of under the category Prevention-but others can be prevention too!).

- Health education

  Traditional type of health promotion (knowledge-attitudes-behaviour-practice).

- Healthy public policy

  Legal, fiscal and regulatory (HIA, European directive).

- Community development

    Radical-individuals setting their own agenda

There are many approaches, as there are several determinants of health (Dahlgren and Whitehead model) that often need to be addressed simultaneously to enable palpable changes. Re: slides on the various wider determinants of health, and the influences on smoking behaviour (slide).

There are many models of Health Promotion. Any Health Promotion model basically gives us a framework for action.

Prevention
There are 4 levels of prevention:
Primordial Prevention
Prevention of factors promoting the emergence of lifestyles, behaviours, exposure patterns which contribute to increased risk of disease.

Primary Prevention
Actions to prevent the onset of disease. To limit exposure to risk factors by individual behaviour change and by actions in the community. Includes health promotion (e.g. health education, prescriptive diets) and specific protection (e.g. vaccination)

Secondary Prevention
To halt progression once the illness is already established. Early detection followed by prompt, effective treatment. Special consideration of asymptomatic individuals.

Tertiary Prevention
Tertiary: rehabilitation of people with established disease to minimise residual disability and complications. Quality of life action even if disease can not be cured.

Approaches to Disease Prevention
There are 2 main approaches to Disease Prevention:
1. *High Risk* - identifying those in special need "targeted rescue operation" (Geoffrey Rose, 1992), then controlling exposure (e.g. reducing house dust mite in the home of asthmatic child) or providing protection against effect of exposure (vaccination).

2. *Population* - begins with recognition that the occurrence of common diseases and exposures reflects the behaviour and circumstances of society as a whole.

Prevention paradox

- Many people exposed to a small risk may generate more disease than the few exposed to a large risk
- Therefore, when many people receive a small benefit the total benefit may be large
- However, individual inconvenience may be high to the many when benefit may only be to a few.

Strengths and weaknesses of high risk approach
Strengths

- Effective (high motivation of individual and physician)
- Efficient (cost-effective use of resources)
- Benefit : risk ratio is favourable
- Appropriate to individual
- Easy to evaluate

Weaknesses

- Palliative and temporary (misses a large amount of disease)
- Risk prediction – not accurate
- Limited potential – misses out on spill over of info
- Hard to change individual behaviours

Strengths and weaknesses of population approach

Strengths

- Equitable (Attributable risk may be high where risk is low if a lot of people are exposed to that low risk)
- Radical
- Large potential for population
- Behaviourally appropriate

Weaknesses

- Small advantage to individual
- Poor motivation of subject
- Poor motivation of physician
- Benefit : risk ratio worrisome

Where can health promotion operate?

- Internationally
- Nationally (government, advertising, media)
- Locally (GP, hospitals, Local Authority, Police, Schools etc)
- Individually (support groups, neighbourhood schemes, communities)

It may impact at the level of:

- The population
- The community
- The individual

Smoking Cessation is a good example of the Health Promotion role of doctors working with individuals

- Smoking cessation guidelines (NICE)
- Motivational interviewing
- Support for cessation
- Prescription of nicotine replacement therapy (NRT) and bupropion (Zyban)
- Referral to specialist services

There is also a broader Health Promotion role that doctors can play - Wider health promotion – Advocacy – E.g. higher taxes, NRT on prescription, ban on tobacco advertising, smoke-free public and work places by:

- writing / speaking to politicians (lobbying)
- letters to the press (media advocacy)
- influencing decision-makers

Does it work? Evaluation through trials

Examples of health promotion programmes

- Seguro Popular – Universal Health Coverage in Mexico
- Healthy City Marikina
- The Sonagachi HIV / AIDS International Partnership

Current Health Promotion programmes in the UK

Key Policy Documents

The Wanless Report

- Wanless 1, 2 and 3
- The Disease Burden
- "Fully Engaged Scenario"
- Focus on prevention and the wider determinants of health
- Cost-effectiveness of actions to improve health and reduce inequalities

Government White paper – Choosing Health

Strategic Review of Health Inequalities in England post-2010 - The Marmot Review

- Give every child the best start in life
- Enable all children, young people, and adults to maximise their capabilities and have control over their lives
- Create fair employment and good work for all
- Ensure healthy standard of living for all
- Create and develop healthy and sustainable places and communities
- Strengthen the role and impact of ill health prevention

Key Public Health Programmes

Smoking Cessation

Alcohol Harm Reduction Strategy

Sexual Health - National Chlamydia Screening Programme

Tackling Teenage Pregnancy

Tackling obesity

Immunisation Programmes

Further Reading

1. *Oxford handbook of Public Health*, Oxford University Press, 2001
2. *The Strategy of Preventive Medicine*, Geoffrey Rose, Oxford University Press, 1993
3. Farmer R, Miller D, Lawrenson R. *Lecture Notes on Epidemiology and Public Health Medicine. (4th ed)* Oxford: Blackwell Science, 1996.

- Ch 12 pp 138-147
- Ch 15 pp 186-195 – Environmental health Ottawa Charter for Health Promotion, WHO 1986 "The fundamental conditions and resources for health are peace, shelter, education, food, income, a stable ecosystem, sustainable resources, social

justice and equity”

4. *Health Promotion Glossary* http://whqlibdoc.who.int/hq/1998/WHO_HPR_HEP_98.1.pdf
5. *Health Promotion: Foundations for Practice*. Jennie Naidoo, Jane Wills
6. *Closing the Gap in a Generation*. Final Report of the Commission on Social Determinants of Health. World Health Organisation, August 2008
7. *Securing Good Health for the Whole Population. The Wanless Report. Derek Wanless, 2004. http://www.hm-treasury.gov.uk/consultations_and_legislation/wanless/consult_wanless04_final.cfm*
8. *Choosing Health: Making healthy choices easier.* Public Health White Paper, Department of Health, November 2004

   http://www.dh.gov.uk/PublicationsAndStatistics/Publications/PublicationsPolicyAndGuidance/PublicationsPolicyAndGuidanceArticle/fs/en?CONTENT_ID=4094550&chk=aN5Cor
9. *Choosing Activity: A physical activity action plan.* Best practice guidance, Department of Health, March 2005
10. *Choosing a better diet: A food and health action plan.* Best practice guidance, Department of Health, March 2005
11. *Fair Society Healthy Lives: Strategic Review of Health Inequalities in England post-2010.* The Marmot Review, 11th February 2010.
12. *Healthcare Promotions UK*: http://www.healthcarepromotions.co.uk/site/index.php?option=com_alphacontent&section=weblinks&Itemid=57. *Updated 19th Sep 2011*
13. *Teenage pregnancy and parenthood: a review of reviews Evidence briefing. Health Development Agency, February 2003*
14. *Lancet 2012; doi:10.1016/S0140-6736(12)61068-X*

Session 6: Tutorial 1

Tools of the trade: understanding and interpreting the findings commonly reported in published papers

Dr Alex Bottle and Dr Mireille B Toledano
(robert.bottle@imperial.ac.uk and m.toledano@imperial.ac.uk)

This tutorial is designed to help you to understand the commonly reported findings you see in papers published in medical journals.

Based on feedback from previous years' students, we have changed the way we teach medical and epidemiological statistics. The focus is now on the interpretation of the statistics rather than their calculation and teaching will be done via a tutorial session rather than in a lecture theatre.

Please note that some of the material in this tutorial will build on what you will learn in sessions 1, 2, 3, 4 & 5 but some of it will not be covered anywhere else in the course. All material taught in this tutorial will be included in your examinations.

The tutorial can be found at the end of the course guide, on page 60

Tutorial groups and venues can be found on page 6

Session 7: Lecture 12

More on Study Design: Systematic Reviews and Meta-analysis

Dr Teresa Norat
(t.norat@imperial.ac.uk)

Learning Objectives

- To understand the need for conducting systematic reviews and meta-analyses.
- To appreciate the potential biases and limitations of systematic reviews and meta-analyses.
- To be able to interpret the findings presented in published systematic reviews and meta-analyses.
- To be able to critically appraise published systematic reviews and meta-analysis.

Why undertake a systematic review?

Because of the high volume of data that need to be considered by practitioners and researchers, it has become impossible for the individual to critically evaluate and synthesise the state current knowledge in many areas. Single studies are often insufficient to universally answer a research question. In order to provide more generalisable conclusions, researchers can conduct a s*ystematic review* of the primary studies on a particular research question.
A systematic review is 'a review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review.'
The advantages of a systematic approach include:

- Transparent process because of the explicit methods in identifying and rejecting studies.
- Meta-analysis, if appropriate, will enhance the precision of estimates of treatment effects.
- Systematic reviews may demonstrate the lack of adequate evidence and thus identify areas where further studies are needed.

What is involved in a systematic review?

There are several stages to undertaking a systematic review:

Stage I
Planning the review - Need to clearly define the research question to be addressed. This question is usually framed around the definition of study participants, intervention (exposure), outcomes and study designs of interest.

Stage II
Identification of research - Requires clearly defined search criteria and a thorough search of all published literature (including exhaustive searches of reference lists, conference proceedings and contact with researchers in the field).

Selection of studies – Inclusion and exclusion criteria should be defined a priori; these are likely to be based on factors such as study design, year, sample size, completeness of information, study quality etc.
Study quality assessment – Study quality can be assessed against recognized or user-defined criteria, usually to establish whether various biases are likely to exist in the in

study (e.g. selection bias, measurement bias, attrition bias/loss to follow-up).

Stage III
Reporting and dissemination – Study details need to be abstracted from each eligible study along with the effect estimate (or details that allow an effect estimate to be calculated). These details need to be tabulated in a meaningful way, including, where appropriate, details of populations, interventions/exposure, outcomes and study design, and a summary of the findings. The last step consists in estimating an overall effect by combining the data, if a metanalysis is deemed appropriate.

What is a meta-analysis?

Meta-analysis refers to 'the use of statistical techniques in a systematic review to integrate the results of included studies'. The studies themselves are the primary units of analysis.
Meta-analyses combine the published estimates of effect from each study to generate a pooled risk estimate. This approach means that:

- More subjects can be included than any single constituent study, producing a more reliable and precise estimate of effect
- Differences (heterogeneity) between published studies can be identified and explored.

However…

- If the studies are too heterogeneous, it may be inappropriate, even misleading to statistically pool the results from separate studies

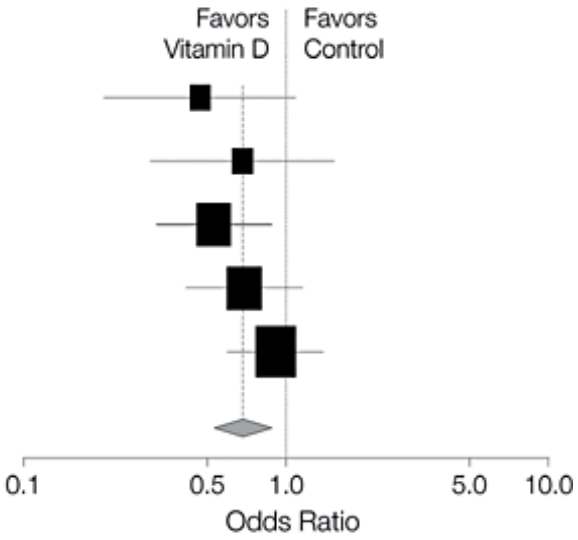
What is involved in a meta-analysis?

As in a systematic review, effect estimates are abstracted (or calculated) from the selected studies; in a meta-analysis, these individual study effect estimates are then pooled to produce a weighted average effect across all studies.

A **Forest plot** is the most common way of presenting the results from a meta-analysis. This is a graphical representation of the results from each study included in a meta-analysis, together with the combined meta-analysis result.

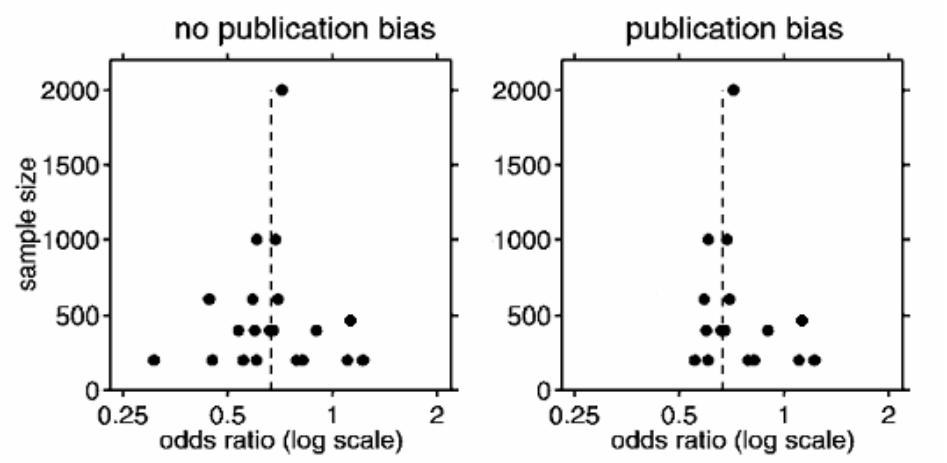
Each study is represented by a box and line – the size of the box corresponds to the weight given to that individual study; the horizontal lines correspond to the 95% confidence interval.

The overall estimate from the meta-analysis is usually shown as a diamond at the bottom of the plot. The centre of the diamond and dashed line corresponds to the summary effect estimate; the width of the diamond represents the confidence interval around this estimate.



Publication bias

Publication bias refers to the greater likelihood of research with statistically significant results to be published in the peer-reviewed literature in comparison to those with null or non-significant results. Failure to include all relevant data in a meta-analysis may mean the effect of an intervention/exposure is over (or under) estimated.
Publication bias in meta-analyses can be explored using **Funnel plots**, which show whether there is a link between study size (or precision) and the effect estimate.



Heterogeneity

Studies that are trying to answer the same question may still differ with respect to the exact population, interventions/exposure, outcomes and designs used. Even where these factors are homogeneous, heterogeneity may still exist because of clinical differences, methodological differences or unknown study characteristics.
Heterogeneity can be explored using **Galbraith (radial) plots**. But remember, if too much heterogeneity exists, it might not be appropriate to pool the studies.

Limitations in conducting systematic reviews

If the methodological quality of studies is inadequate then the findings of reviews of this material may also be compromised.
Publication bias can distort findings because studies with statistically significant results are more likely to get published.

Further reading

Centre for Reviews and Dissemination. *Undertaking Systematic Reviews of Research on Effectiveness*. CRD Report Number 4, March 2001 (available from http://www.york.ac.uk/inst/crd/report4.htm)

Egger, M, Davey Smith, G, Altman, eds. DG. *Systematic Reviews in Health Care*, BMJ Publishing Group, 2001.

Greenhalgh, T. *How to read a paper: Papers that summarise other papers (systematic reviews and meta-analyses).* BMJ 1997;315:672-675.

Khan, KS, Kunz, R, Kleijnen, J & Antes, G. *Systematic Reviews to support Evidence-based Medicine.* The Royal Soceity of Medicine Press Ltd. 2003.

The Cochrane Collaboration: http://www.cochrane.org

Session 7: Lecture 13

Introduction to critical appraisal and how to read published papers

Dr. Claire Robertson
(C.Robertson@westminster.ac.uk)

Learning Objectives

At the end of this session you will be able to:

- Search for evidence from published medical journals on a specified research topic
- Review and critically appraise the evidence presented in a scientific paper
- Present critical appraisal findings to lecturers and peers

Background
Critical appraisal is the process of systematically examining research evidence to assess its validity, results and relevance before using it to inform a decision. It constitutes an essential part of evidence-based clinical practice, allowing us to make sense of research evidence and begin to close the gap between research and practice. All graduates are expected to have demonstrated an ability to use both scholarly reviews and primary sources of data including published papers and/or original materials (The Quality Assurance Agency for Higher Education, 2008). It is important to bear in mind when doing this that the quality of study designs is never certain, despite publication in peer-reviewed journals.

Development of critical appraisal skills enables you to find, and make sense of, research evidence, and to put newly gained knowledge from research appraisals into practice. Three phases are involved in critical appraisal: finding, appraising, and acting on research evidence.

1. Finding research evidence
Unless you use a systematic method to find your references, you will waste a lot of time, and will potentially miss a great deal of important articles. Keeping in mind that this summary cannot synopsize all aspects of the ideal method to use in a systematic literature search, it covers some of the key points to bear in mind when searching for references.

Defining the topic: Start by writing a clear question which you aim to answer using the papers found in your literature search. Once this is done, consider what 'key words' might be used by authors publishing in this area, and use these to search for appropriate papers. It is important to remember that even starting with a clear definition you may find more papers than you can cope with. To deal with this, refine your search (e.g., look for articles published within the last 10 years and in English only) to generate a more manageable or focused set of results. In contrast, if only 5 articles are found, you may want to consider broadening your search terms (e.g., the population of interest) to identify more articles. If you decide to refine your search terms by study design, keep the research question you are attempting to answer in mind. A randomized controlled trial (RCT) for example cannot tell you about how easy patients find adherence to a set of guidelines. It could however help you to identify why a particular guideline is relevant to your patients care by assessing its effects within controlled situations, and therefore including such papers may be important.

Identifying sources of information: Published information is often the only source of information utilised, but additional sources are also available, including:

- Electronic search engines e.g., Medline, PubMed, Sciencedirect (NB, this is not a complete search engine as it includes only Elsevier journal articles). http://www.imperial.ac.uk/library/eresources/M1.html lists the college's available electronic resources. Library staff can help structure your searches;
- Acknowledged experts;
- Practitioners;
- Theses, conference papers etc.

It is important to bear in mind that while the internet contains a great deal of factually correct and useful information, it also contains a great deal of unsubstantiated and incorrect information. As a general rule therefore, it is preferable to avoid using internet-based information, unless this is obtained from (and therefore can be referenced to) published sources, academic or government institutions, or charities for example. The Department of Health launched 'The Information Standard' in 2008 to enable the public to identify health and social care information that is accurate and can be trusted. Their identifying logo (see below) is found for example on Bupa, Breast Cancer Care and several NHS Foundation trust websites. Full lists of certified websites are listed at: http://www.theinformationstandard.org/

Keeping records: This is perhaps the most important adjunct to your literature review. Accurate bibliographic details, search histories, critique details, key information from papers etc will help you find things again quickly. Reference manager and Endnote are useful electronic systems; index cards are a more classic format.

2. Appraising research evidence

A number of systems can be used to organize your review of research evidence. These can either generic systems or systems specifically focused on appraising evidence from specific study designs.

Below is one example of a generic checklist you can consider when reviewing any evidence found regardless of study design. This is illustrated using 9 separate questions which you should ask:

A. Generic appraisal checklists

a) The Question?

What is the question the researchers are trying to answer? Is there a hypothesis? Is the question relevant?

Is the research original? Consider…

- The size of the study – is it bigger, continued for longer, or more substantial than previous studies?
- Are the methods used any more rigorous than those published previously?

- Will the numerical results add significantly to any subsequent meta-analysis?
- Is the population studied different in any way (e.g., ethnicity, age, sex)?
- Is the clinical issue being addressed important?

b) Design

Is it cross-sectional, cohort, case-control, ecological, RCT? Refer to the hierarchy of studies. Is the study design appropriate?

A hierarchy of evidence is frequently used to summarise the strength of data traditionally associated with each study design (the first being the strongest example). It is important to remember however that the design of each study is also crucial.

- Systematic reviews of randomized controlled trials (RCT's)
- Individual RCTs with narrow confidence intervals
- All or none-case series (when all patients died before a new therapy was introduced, yet following its introduction, all patients receiving the treatment survived)
- Systematic review of cohort studies
- Individual cohort study or RCT with <80% follow up
- Outcomes research: Ecological studies
- Systematic review of case control studies
- Individual case control study
- Case series
- Expert opinion

c) Population?

Sample size

- Has a power calculation been conducted?
- Are results generalisable to other populations
- It is important to bear in mind that findings from a clinical trial conducted in participants diagnosed with a disease may not be the same as those seen in those without the disease. In addition, ethnic groups or those exposed to cigarette smoke (for example) may produce different results.

How were the subjects recruited?

Recruitment by newspapers may introduce bias (as only those readers motivated to respond will be studied). It is better to invite all people in your target population (e.g., secondary school teachers in London, nurses employed at grade H at St Mary's hospital), or a random sample of these, to minimise this error.

Inclusion criteria

These criteria will essentially define the population to which results can be extrapolated, and can benefit the smooth running of the trial (e.g., by including only literate, English-speaking participants).

Exclusion criteria

This definition will refine your target population and remove avoidable sources of bias. For example, excluding patients with a co-existing illness, pregnant or lactating women, or men aged 40 and over.

Were the subjects studied in "real life" circumstances?

If we ask people to consume a 'study diet' in a clinical setting, we have insufficient evidence to prove whether in a free-living situation, with access to other lifestyle choices (e.g., foods, exercise, smoking, alcohol), the results would hold. Additional factors, for example, increased contact with physician or use of new equipment could similarly affect the application of study findings.

d) Methods

What specific intervention was being considered and what was it being compared with?

A report of "We measured how often GPs ask patients whether or not they smoke" could be improved by stating "We looked in the patients' medical records and counted how many had had their smoking status recorded." However, this assumes that medical records are 100% complete and accurate therefore is not unflawed.

What outcome was measured and how?

And how does it relate to the disease mortality and morbidity rates? Measurement of symptoms, pain, psychological measures are more difficult to estimate than biochemical tests, you should always check whether the methods are validated and that its inclusion is necessary.

Duration of follow-up

Has the study continued for long enough to detect the effect of the intervention?

e) Analysis

- Have the appropriate statistical tests been used?
- Did the authors take into account chance and adjust for confounding?

f) Confounders

- Is there the possibility of confounders which have not been adjusted for?

g) Bias

Measurement/selection?

What techniques were used, and were these appropriate? A systematic bias could be introduced if a thermometer was incorrectly calibrated 3° lower than the actual temperature. Similarly, if a person chose to report consumption of semi-skimmed milk, when whole milk was in fact consumed, this would constitute a systematic error. This differs from random error.

Randomised controlled trials

Avoid systematic bias in a RCT by selecting participants from a particular (defined) population and allocating them randomly to different groups.

Non-randomised controlled studies (cohort & case control studies)

It is almost impossible to identify two groups of subjects with the same age/gender mix, socio-economic status, presence of co-existing illness etc – therefore adjustment must be made for these differences between populations using appropriate statistical methodologies. Considering the effects of alcohol intakes on health outcomes (cohort study) between 2 population groups for example may be affected dependent on how data collected from non- and ex-drinkers is handled. In case control studies, the definition of a 'case' and a 'control' should be crystal clear and follow an evidence-based rationale to ensure the interpretation of subsequent findings is clear and safe.

Completeness of follow-up

Is any assessment made of those who dropped out of the study? If this was due to adverse effects associated with drugs being studied, are the results accurate? Ignoring drop-outs will bias results.

Was the assessment blind (double blind)?
Several research study designs require that the participants and/or the researchers are unaware of whether or not the participant has received the placebo or trial drug. If, for example, patients were applying cream to a wound and nurses were addressing their improvement, expectation could potentially bias the recorded results.

h) Ethics
Is the study ethical? Is informed consent obtained? Does other known research indicate that there is a reason why the study design should not have continued? etc.

i) Interpretation
- Do the authors interpret their collected data correctly?
- Do they make a causal inference (remember Bradford-Hill)?

B. Specific appraisal checklists

Depending on the epidemiological study design employed there are specific checklists that can help you evaluate the research undertaken. These focus on the specific strengths and weaknesses of each design. They are:

- STARD
- STROBE
- MOOSE
- CONSORT Checklist
- CONSORT Flowchart
- QUORUM-PRISMA

Each of these specific checklists can be found at the back of this handbook, together with the material for Tutorial 2.

3. Presenting critical appraisal findings to lecturers and peers

It is important that you summarise the paper first to get your own ideas about it clear. A sentence for each of the following will often suffice

- Why did they do it?
- What did they do?
- What did they find?
- What did they conclude?
- In your opinion, was the study conducted well?

If you are reviewing a number of papers in one area, it is often worth doing this within a series of tables. This will be of great use if you wish to conclude giving a 'balance of evidence' overview: by considering the number of papers you have found (i.e., number of rows of information in your table), you can quickly refer to the number of studies which identified a direct, inverse and no association – and using your critique notes for each study design, use this to rationalize why your expected relationship may not have been evidenced.

Further reading

The following two texts direct you to good places to start additional reading – remember that your reading should not stop here however!

- Trisha Greenhalgh (2001) How to Read a Paper: The Basics of Evidence Based Medicine (2nd edition). BMJ Books; London.

- Bad Science. By Ben Goldacre. *Particularly the chapter on the pharmaceutical companies.* Fourth Estate Ltd 2008. ISBN-10: 0007240198

Session 8: Lecture 14

Understanding and appraising evidence: Horizon video and questions

Dr Susan Hodgson
(susan.hodgson@imperial.ac.uk)

Video: The Valley of Life or Death (Horizon, BBC)
http://www.bbc.co.uk/science/horizon/valley_hiv.shtml

“At the heart of the AIDS epidemic in Africa, there is a deadly mystery that has puzzled scientists for years. There are groups of people who are four times less likely to get HIV than other people, sometimes living just yards away, across a single valley - people with apparently similar behaviour and lifestyle. Scientists realised that if they could understand why these people are so much less vulnerable to the HIV virus, it might lead to an answer that could save millions of lives. And after 15 years of detective work it turns out there may be a remarkably simple answer: the high risk areas for HIV coincide with tribes who are uncircumcised. In Africa, it seems a man is much more likely to get HIV if he is uncircumcised.

In Kaoma, Western Zambia, a young boy is on his way to the sacred Mukondaa - the tribal circumcision ground. Around him the tribal elders are gathered, dressed in their ceremonial garb, and vivid masks. But the young boy himself is an outsider, not from this tribe, and none of his relatives or ancestors have ever been circumcised. In fact, his parents are only prepared to break the taboo of their own tribe because they believe that circumcision could save his life by protecting him from AIDS. At first sight this belief seems like the kind of superstition to which desperate people often turn in times of plague. But now there is scientific evidence that suggests these people could well be right.

There have now been twenty seven statistical studies that show a big difference in HIV infection between circumcised and uncircumcised men. For example, among the uncircumcised people of Kisumu in Western Kenya, a man is three times as likely to get AIDS as his circumcised neighbours. Among truck drivers in Mombassa the difference is four-fold.

‘Horizon’ travels across Africa, tracing the work of scientists who have unearthed the statistical data behind this correlation. At the same time microbiologists have been battling to understand the complex and insidious virus, and their work indicates that the foreskin may be a key entry point for HIV. The logical conclusion for these scientists is that if you remove the foreskin, you begin to protect the man. No-one believes that circumcision can protect completely - the evidence so far only indicates that it reduces the risk of infection by HIV, and then only during heterosexual sex. Unquestionably, condoms are still the best protection. But in the many countries where the use of condoms is minimal, it seems that circumcision might help to reduce the spread of AIDS.

In the absence of a vaccine for AIDS, and the lack of condom use in the developing world, should governments think the unthinkable and encourage the circumcision of young boys in non-circumcising tribes as a public policy? Opposing this idea are the voices of tribal elders who are loath to change tribal traditions that have existed for generations, and a fierce Western anti-circumcision lobby which believes that circumcision is a form of mutilation and violates basic human rights.

Session objectives
At the end of this session you will be able to

- Define prevalence
- Describe methods for measuring prevalence
- Calculate relative risk from a simple example
- Discuss interpretations of an unexpected association.

Background

Ritualistic circumcision has been carried out in West Africa for over 5000 years, and in the Middle East for at least 3000 years. In the USA and Canada, circumcision appeared as part of the medical culture during the late 19th and early 20th century, and by the early 1970s about 80% of US newborn boys were being circumcised.
There is little evidence of any health benefit, and the American Academy of Paediatrics opposed routine neonatal circumcision in 1971, and the rate has subsequently declined.
About 25% of men in the world are circumcised, largely in the USA, Canada, Middle East and Asian countries with Muslim populations, and large portions of Africa.

During this session, you will see part of a video on HIV infection in Africa exploring the link between male circumcision and HIV risk. We will consider some of the epidemiological data on this association.

Study design exercise

Part 1

The video Valley of Death reports the association between HIV and lack of circumcision in men. For example, the prevalence of HIV in tribe A was 7% compared with 21% in tribe B.

(a) What is prevalence?

(b) How do you think the prevalence of HIV was measured in the two tribes?

(c) What are the problems of the different methods of determining prevalence?

Part 2

In the video the association between HIV and lack of male circumcision was shown in a number of studies. Frank Plummer reported a study in Nairobi where HIV negative men who reported recent contact with HIV infected women (prostitutes) were followed up to see how many of them acquired the infection. (Cameron et al 1989). This is called a cohort study.
This table shows the basic results

		After 3 months		
		HIV+	HIV-ve	Total
Circumcised	no	18	61	79
	yes	6	208	214
		24	269	293

(a) What is the incidence in the non-circumcised men?

(b) What is the incidence in the circumcised men?

(c) What is the relative risk?

(d) What does this mean?

Part 3

What other evidence was presented in the video to show the association?

Part 4

It is possible, and some would say biologically plausible, that circumcision of men protects against HIV infection. However, what other explanations could there be for the observed association?

Session 8: Tutorial 2

Critical Appraisal of Medical Evidence

Dr Mireille B. Toledano and Dr Claire Robertson
(m.toledano@imperial.ac.uk and C.Robertson@westminster.ac.uk)

This tutorial is designed to help you understand how to read and interpret the evidence presented in papers published in medical journals.

Based on feedback from previous years' students, we have changed the way we teach critical appraisal of medical evidence. Students felt that they would benefit from small group tutorial sessions where a tutor would be available to help clarify important points, answer questions, and facilitate discussion. Please note that this tutorial will be an opportunity to *consolidate* and *apply* the material you will learn in sessions 1-8 (and in particular Lectures 13-14) and test how well you have understood the concepts covered. **All material taught in this tutorial will be included in your examinations.**

Following on from Lecture 13 (Introduction to critical appraisal of medical evidence), this tutorial is focused upon two published papers investigating the effects of vitamin supplements on cardiovascular disease and mortality (both of these papers are provided for you at the back of this handbook):

Paper 1: Pocobelli G, Peters U, Kristal AR, White E. Use of supplements of multivitamins, vitamin C, and vitamin E in relation to mortality. *Am J Epidemiol 2009:170; 472-483.*

Paper 2: Lee IM, Cook NR, Gaziano JM, Gordon D, Ridker PM, Manson JE, Hennekens CH, Buring JE. Vitamin E in the primary prevention of cardiovascular disease and cancer: the Women's Health Study: a randomized controlled trial. *JAMA 2005: 294(1); 56-65*

You should read these two papers before the tutorial so that you have sufficient time to work through the tutorial in your groups during the timetabled session with your tutor.

The full tutorial can be found at the end of the course guide, on page 70.

Tutorial groups and venues can be found on page 6

Session 9: Lecture 15

International Health I
Part 1: Poverty, Health and Development

Mr Mike Rowson
(m.rowson@ich.ucl.ac.uk)

Learning Objectives

By the end of this session you should be able to:

- describe the extent of health and income inequalities worldwide
- describe poverty and child mortality rates in different parts of the world
- describe the relationship between GDP per capita and child mortality rates across developed and developing countries
- understand some of the key factors that might explain why some countries with similar incomes achieve variant child health outcomes

---

Session description

Inequalities worldwide in health and income are extreme. Commonsense suggests a relationship between income and health outcomes, but statistics at the national level show some interesting variations. Using Gapminder, we will examine factors that either mediate the income and health relationship or which by themselves explain the variance in health outcomes between nations more fully.

Further reading

Cutler D, Deaton A, Lleras-Muney A (2006). The determinants of mortality. *Journal of Economic Perspectives* 20 (3):97-120.

Gapminder. http://www.gapminder.org/downloads/flash-presentations/human-development-trends-2005/

Hulme D (2010). Lessons fom the making of the MDGs: human development meets results-based management in an unfair world. *IDS Bulletin* 41(1):15-25.

Parker M et al. (2001). Diseases of poverty. In Allen T and Thomas A (eds). *Poverty and Development into the 21st Century*. Oxford, OUP.

Sen A (1998). Mortality as an indicator of economic success and failure. *The Economic Journal* 108:1-25.

United Nations Development Programme (2010). *Human Development Report 2010*. Oxford, OUP. Chapter 2
http://hdr.undp.org/en/media/HDR_2010_EN_Chapter2_reprint.pdf

Vandermoortele J (2009). The MDG conundrum: meeting the targets without missing the point. *Development Policy Review* 27 (4):355-371.

International Health I
Part 2: Globalisation and Health Worker Migration

Mr Mike Rowson
(m.rowson@ich.ucl.ac.uk)

Learning Objectives

By the end of this session you will be able to:

- outline the reasons for health worker migration from poorer to richer countries
- assess how feasible it is to prevent health worker migration in the context of globalised labour markets of health professionals
- evaluate some of the proposed solutions for problems exacerbated by health worker migration.

Session description

During the first part of this decade, health worker migration or "brain drain" received a large amount of publicity as a key issue in the health crisis affecting poorer countries. Health worker migration is one aspect of globalisation – in this case the creation of integrated labour markets for health professionals. But how far can this type of globalisation be prevented, given the strong push and pull factors that cause migration and in the light of technological development and the subsequent "death of distance" that comes with it? If globalisation creates greater health inequality, is it best to stop the globalisation (prevent workers from migrating) or to look for other ways of creating "incentives to stay" and thus address the profound inequalities in health care that exist between rich and poor countries. The session will debate these controversial issues.

Further reading

Buchan J (2010). Can the WHO code on international recruitment succeed? *British Medical Journa*l 340:791-793.

Bueno de Mesquita J and Gordon M (2005). The International Migration of Health Workers: A Human Rights Perspective. London, Medact. http://www.medact.org/content/Skills%20drain/Bueno%20de%20Mesquita%20and%20Gordon.pdf

Hongoro C and McPake B (2004). How to bridge the gap in human resources for health. Lancet 364:1451-1456.

Mackintosh M, Mensah K, Henry L, Rowson M (2006). Aid, restitution and international fiscal redistribution in health care. Journal of International Development 18:757-770.

Mensah K, Mackintosh M, Henry L (2005). The Skills Drain of Health Professionals from Developing Countries: A Framework for Policy Formulation. London, Medact. http://www.medact.org/content/Skills%20drain/Mensah%20et%20al.%202005.pdf

World Health Organization (2006). *World Health Report 2006 – working together for health*. Geneva, WHO. http://www.who.int/whr/2006/en/

WHO (2010). WHO code of practice on the international recruitment of health personnel. Geneva,
WHO. Available at http://www.who.int/hrh/migration/code/practice/en/index.html

Session 9: Lecture 16

International Health II
The integration of global health partnerships to combat the propagation of waterborne infectious diseases in Africa

Professor Alan Fenwick
(a.fenwick@imperial.ac.uk)

Learning objectives

- To demonstrate an understanding of global health issues with regards to waterborne infectious diseases

---

Summary

There are a number of infectious and parasitic diseases which are the scourges of Africa. They affect many millions of people and cause much disability and suffering without causing enough deaths to be considered as acute diseases like malaria and HIV/AIDS. Because they affect the poorest of the poor they tend to be neglected and left untreated, and in fact we have placed them all together under the best name we could come up with - "Neglected Tropical Diseases". These diseases include intestinal worms (hookworm and ascaris), schistosomiasis, river blindness, elephantiasis, leprosy, trachoma and sleeping sickness.

For some the strategy used is diagnosis and treatment (eg leprosy and sleeping sickness). For the others, they are susceptible to 4 drugs (Mectizan, Albendazole, Zithromax and Praziquantel), and in fact an annual mass treatment with these drugs - given in pairs six months apart should eliminate the morbidity due to these diseases in less than 8 years. The drugs are for the most part donated as part of the pharmaceutical industries humanitarian donation programme, but unfortunately a condition of the donation is that countries need to identify a source of the funds needed for heath education, training, distribution, monitoring and evaluation. While this amounts to only about 25 pence per person per year, because the treatment target numbers are so great (say 500 million in Africa) some £125 million pounds per year is needed to complete the job. Not huge in terms of total health spending, but for a poor country it is enough to stretch resources.

Since the Millenium Development Goals were announced, it has been realised that control of these NTD's would contribute to at least 6 of the 8 MDG's, and so Global Health Partnerships including disease control programmes and the pharmaceutical industry, have come together to try and combat these diseases using an integrated strategy.. I describe these infections and will report the latest situation with regard to control, not only by the Imperial College Schistosomiasis Control Initiative SCI), but also by a USAID supported NTD control project, and by the Global Network for the Neglected Tropical Disease Control (GNNTDC). Please visit some useful websites before attending the lecture www.schisto.org www.gnntdc.org www.trachoma.org and www.filariasis.org

Session 10: Lecture 17

Disease Prevention II: Screening

Dr. Bhargavi Rao
(bhargavi.rao@imperial.ac.uk)

Learning Objectives

- To understand the principles and practice of screening
- To be able to define validity for screening tests and calculate specificity, sensitivity and predictive value
- To understand the criteria for screening programmes

Definition

Screening is the practice of investigating apparently healthy individuals with the object of detecting unrecognised disease or its precursors in order that measures can be taken to prevent or delay the development of disease or improve prognosis

Purpose of screening

Screening is carried out where the detection of disease at an early stage leads to improved **prognosis** (see glossary). If earlier detection does not offer any hope of improved outcome then screening is generally not indicated. For example, earlier detection of breast cancer allows treatment (surgery, radiotherapy and chemotherapy) that can reduce mortality (leading to increased survival).

Screening may also be used for **risk factors**, i.e. to identify people at increased risk of developing disease where interventions will reduce that risk (for example screening for high blood cholesterol levels or high blood pressure, and then offering lifestyle advice and /or drug therapy to reduce the risk of cardiovascular disease).

Screening may also be used to identify people with **infectious disease** where treatment or other control measures will improve the outcome for the individual (e.g. chlamydia screening), or prevent ongoing transmission to others (e.g. screening food handlers for salmonella, health workers for hepatitis B).

Limitations

The concept of screening is appealing. However, by definition screening tests are carried out on apparently healthy individuals and it is always possible that screening may, inadvertently, do more harm than good. This could include false alarms, inducing anxiety, and the treatment of early disease which would not otherwise have become a problem. When considering population screening programmes the benefits and harms must be carefully assessed, and the benefits should always outweigh the harms.

For example, one study of breast cancer screening showed that for every 50,000 screens carried out, 2820 women would be found to have “abnormal” results requiring further investigation. Only 129 of these turned out to be invasive cancer. While mortality in the population was reduced, there are also considerable costs associated with the identification of women with “abnormal results” who face further investigation and considerable anxiety.

Screening tests

A screening test is not the same as a diagnostic test. The former is usually cheap and simple, and aims to identify people with precursors of the condition or at high risk of the

condition. Further **diagnostic tests** are then done to confirm diagnosis.

The **validity** of any test is its ability to distinguish between subjects with the condition and those without.

To assess the validity of a screening test the true disease status of the individuals must be known, usually through a definitive test which is referred to as the **gold standard**.

Validity is described in terms of sensitivity and specificity of the test (see the figure). An additional test parameter is the predictive value. This is particularly useful in clinical practice.

Figure 1.

		Disease status*		
		Diseased	Non-diseased	
Test result	Positive	a	b	a+b
	Negative	c	d	c+d
		a+c	b+d	

*according to gold standard

The **sensitivity** is the ability of the test to correctly identify people with the disease

sensitivity = a ÷ (a+c)

The **specificity** is the ability of the test to correctly identify people without the disease

specificity = d ÷ (b+d)

The **positive predictive value (PPV)** is the likelihood that a patient with a positive test result that will actually have the disease

positive predictive value = a ÷ (a+b)

The **negative predictive value (NPV)** is the likelihood that a patient with a negative test result that will not have the disease

negative predictive value = d ÷ (c+d)

The predictive value of a test is dependent on the sensitivity and specificity AND the prevalence of the condition in the population (see example at the end).

Approaches to screening

Screening can either involve the whole population **(mass)**, or selected groups who are anticipated to have an increased prevalence of the condition **(targeted)**. In either of these there may be a **systematic** programme where people are called for screening (e.g. cervical cancer, breast cancer) or an **opportunistic** programme when a person presents to the doctor for some other reason and they are offered a test (e.g. Chlamydia screening in young people, blood pressure screening in older people).

Major screening programmes in the UK

Antenatal screening: syphilis, HIV, hepatitis B, rubella, chromosome abnormalities, foetal growth etc. Some of these are offered to all pregnant women, others are based

on risk assessments.

Neonatal and childhood: Newborn babies are screened for phenylketonuria, hypothyroidism, haemoglobinopathies and sickle cell disease (in some geographical areas where these conditions are more common). Babies are also checked for congenital hip dislocation. Routine checks in later childhood screen for problems with hearing and development.

Cancers There are systematic programmes for **breast** cancer and **cervical** cancer in women. A screening programme for **bowel** cancer has started 2006) for all men and women aged 60 – 69. There is no systematic screening programme for **prostate** cancer at the moment, although this is under review.

Infections A new national opportunistic screening programme for chlamdyia in young people (under 25) is currently being rolled out across the country. People attending sexual health services are offered screening for HIV. Hepatitis B screening is mandatory for health care workers.

Cardiovascular disease Targeted and opportunistic screening is carried out for blood pressure, high cholesterol, diabetes in primary care.

Criteria for Screening (based on WHO criteria)

Disease	important health problem well recognized pre-clinical stage natural history understood long period between first signs and overt disease
Diagnostic test	valid (sensitive and specific) simple and cheap safe and acceptable Reliable
Diagnosis and treatment	facilities are adequate effective, acceptable and safe treatment available cost effective Sustainable

Evaluating screening programmes

Even after a disease is determined to be appropriate for screening and a valid test becomes available, it does not necessarily follow that a widespread screening programme should be implemented. Evaluating of a potential screening programme involves consideration of three main issues:

1. Feasibility

Feasibility will depend on how easy it is to organise the population to attend for screening, whether the screening test is acceptable, whether facilities and resources exist to carry not the necessary diagnostic tests following screening.

2. Effectiveness

Effectiveness is evaluated by measuring the extent to which implementing a screening programme affects the subsequent outcomes. This is difficult to measure because of a number of biases that affect most of the study designs used:

> **Selection bias** exists as people who participate in screening programmes often differ from those who do not.
>
> **Lead time bias** exists because screening identifies disease that would otherwise be identified at a later stage. This may result in an apparent improvement in the length of survival due to screening which is really due to the earlier date of diagnosis
>
> **Length bias exists** as some conditions may be slower in developing to a health threatening stage, that is, they have a longer preclinical stage. This means they are more likely to be detected at that stage but they may also have a more favourable prognosis leading to the false conclusion that screening is beneficial in lengthening the lives of those found positive.

3. Cost

The cost of screening programmes is important. Resources for health care are limited and there are many competing demands for available money, health care professionals and facilities. The relative cost-effectiveness of a screening programme compared with other forms of health care should therefore be considered. Costs relate not just to the implementation of the screening programme but also to the further diagnostic tests and the subsequent cost of treatment. On the other hand, in the absence of screening, costs will be incurred by the treatment of patients in more advances stages of disease.

4. Ethics of screening

A screening test is a medical intervention that is done to a person who is not ill and usually to someone who has not initiated the request for the test. For this reason the ethics of carrying out screening must be carefully considered.

- For the individual the screening test can do harm as well as giving benefit
- There may be a risk attached to the screening test or subsequent diagnostic test
- A false positive result can cause unnecessary anxiety
- There may be other unplanned effects of a positive test
- A false negative result will give false reassurance

Glossary

Prognosis - is the outcome of an illness, including duration of disease, mortality and morbidity.

Sensitivity - the ability of a test to correctly identify people with the disease

Specificity - the ability of a test to correctly identify people without the disease

Positive predictive value - the proportion of positive test results that actually have the disease

Negative predictive value - the proportion of negative test results that do not have the disease

Gold standard - a recognised way of determining who really has the disease

Prevalence - the proportion of people in a population with a disease

Further reading/links

www.cancerscreening.nhs.uk An overview of cancer screening programmes in the UK
www.nsc.nhs.uk An overview of current issues related to screening in the UK

Tutorial 1 - Tools of the trade: understanding and interpreting the findings commonly reported in papers

Dr Alex Bottle, Dr Mireille B. Toledano

Instructions:

This tutorial is designed to help you to understand the commonly reported findings you see in papers published in medical journals.

Based on feedback from previous years' students, we have changed the way we teach medical and epidemiological statistics. The focus is now on the interpretation of the statistics rather than their calculation and teaching will be done via a tutorial session rather than in a lecture theatre. Please note that some of the material in this tutorial will build on what you will learn in lectures 5, 6, 7, 8, 9, and 10, but some of it will not be covered anywhere else in the course. **All material taught in this tutorial will be included in your examinations.**

The introductory text describes two worked examples. These examples have been provided to teach you core concepts and to help put into context what you have already learnt; there is also a glossary at the end of the tutorial to define the key terms you will need to know, these terms are italicised in the text. You should read these worked examples before the tutorial, so that you have sufficient time to work through the questions provided during the timetabled session with your tutor.

Learning outcomes:

- Be able to understand the concept of sampling and sampling variation
- Be able to understand that from a *sample*, estimates of the true underlying *risk* in a population can be calculated.
- Be able to define and interpret a *P value* and a *confidence interval*
- Be able to explain the role of statistical hypothesis testing and *confidence intervals* when dealing with chance
- To know the difference between probability and odds and be able to interpret and explain measures of association (*relative risk*, *attributable risk*, *odds ratio*) from simple examples
- Define *confounding* and understand the problems associated with it. Be able to list some methods for dealing with *confounding* (including *stratification*, ***standardisation*** and ***regression***).

Suggested further reading:

Martin Bland (2000) ***An introduction to medical statistics***. Oxford University Press.

Worked example 1 (sampling, P values and confidence intervals)

What is the role of statistics in medicine? Discuss!

1.1 Sampling – estimating prevalence of disease or risk factors

A Primary Care Trust (PCT) wants to estimate the *prevalence* of smoking among their 100,000 residents. What does prevalence mean? How would they do this?
Suppose they surveyed a random *sample* of people – why take a random sample?
Suppose they asked 100 people if they smoked and found that 28 did. If they then asked another 100, would they also find that 28 of them smoked? Why might they not?

If they kept sampling sets of 100 people and plotted the percentage of smokers (prevalence of smoking) in each sample, we would expect to see a *normal distribution* (see glossary), with most sample estimates centred around the true population percentage.

1.2 Confidence intervals and P values – assessing the role of chance

The PCT's estimate of their population's smoking prevalence is 28% from their sample, but there will be some uncertainty around this estimate. We express this uncertainty using a 95% *confidence interval* (95% CI) around the estimate, e.g. 19% to 37%. This means that if we repeated the sampling 100 times, we would expect the true prevalence of smoking in the PCT to fall within the CI in 95 of the 100 samples.

Suppose the PCT wanted to lower this prevalence; they could implement a smoking reduction campaign and then see if it worked by comparing their first estimated prevalence with an estimate after the campaign. They took two random samples, the first finding that 28% smoked as above, and the second finding that 21% smoked. Can we therefore say for certain that the campaign has worked and cut the prevalence by 28-21=7%?
Why not?
We want to know whether the difference of 7% could simply be due to chance (sampling error) or is a real difference in prevalence. This is done statistically by setting up a *null hypothesis* of no difference and looking for evidence to disprove it: what is the likelihood that our two samples were 28% and 21% if the two true underlying prevalences were the same? We then choose the appropriate statistical test (e.g. chi-squared test to compare the two proportions) to get this likelihood, which is the *P value*. The lower the P value, the less likely that our estimated difference is a chance finding. Suppose the P value was 0.014. Convention has it that if $P<0.05$ (and this is an arbitrary cut-off!) then we can reject the null hypothesis and conclude that the smoking prevalence fell after the campaign. Such a result is called statistically significant.

THE PROCEDURE:

1. Set up a null hypothesis (e.g. difference in prevalence between the two groups is zero)
2. Choose an appropriate statistical test
3. Inspect the results (estimated measure of association – or, in this case, estimated difference in prevalences – plus its CI and P value) for evidence of real difference: can we reject the null hypothesis?

Are statistically significant results more or less likely with small sample size than with large sample sizes? Why (the answer is to do with the nature of the sample rather than statistics)?

Worked example 2 (measures of association)

The main aim of epidemiological research is to investigate the association between exposure to a risk factor (e.g. smoking) and the occurrence of disease (e.g. lung cancer). We compare the incidence in a group of people exposed to the risk factor with a group who were not exposed. Suppose the incidence in one group is higher than in the other – what are the two different ways of stating this? If Joe is 36 and John is 18, how could we say by how much Joe is the elder?

2.1 Ratio measures: relative risk and odds ratio

Two key concepts: risk and odds. What is the difference?

Suppose you wanted to look at possible risk factors for lung cancer: smoking and occupational exposure. How might you select your population sample to do this?

COHORT STUDY: Malarcher et al (2000) took a large group of US males, some smoke and some never have done, and followed them up over time. They measured the *rates* of lung cancer in the two groups. They set up the null hypothesis of equal rates and calculated the *relative risk*: 27 for smokers compared with those who have never smoked (95% CI 19 to 38). Can we reject the null hypothesis?

The interpretation of a relative risk is straightforward: if you smoke, you are 27 times more likely to die from lung cancer than if you don't smoke.

CASE-CONTROL STUDY: Richiardi et al (2005) took a group of people with lung cancer (the *cases*) and another without lung cancer (the *controls*) and asked each about their occupation (whether they were dockers or freight handlers). Their occupation is the exposure here. Richiardi measured the odds of exposure (odds of working as a docker or freight handler) in the cases and then in the controls. They set up the null hypothesis of equal odds and calculated the *odds ratio*: 1.5 for those with lung cancer compared with those without (95% CI 1.1 to 2.1). Can we reject the null hypothesis?

This odds ratio means that someone with lung cancer is 1.5 times more likely to have worked as a docker or freight handler than someone who doesn't have lung cancer. Notice that it compares the exposure in the two groups – it does not compare the disease rates in the two groups, which the relative risk does. The odds ratio is an estimate of the relative risk, and it is usually more useful (and easier!) to interpret an odds ratio to mean that if you work as a docker or freight handler you are 1.5 times more likely to get lung cancer than if you work in a different occupation. See the glossary for an explanation of the relationship between relative risk and odds ratio and on why case-control studies can only provide us with the latter.

2.2 Difference measure: attributable risk (or attributable fraction)

The attributable risk for lung cancer in smokers is the rate of lung cancer amongst smokers minus the rate of lung cancer amongst non-smokers (i.e. the risk difference). It gives an indication of how many extra cases for which the exposure is responsible, making the important assumption that the relation between the exposure and the disease is causal (i.e. not explained by other confounding factors – see below). The attributable risk and related measures are typically used to help guide policymakers in planning public health interventions.

2.3 Confounding – and controlling for it

How can we prove that an exposure causes a disease, rather than is merely associated with higher rates of that disease? We try to eliminate (i.e. control or adjust for) the effects of confounders. Confounders are associated with both the exposure of interest and the outcome of interest (e.g. developing a disease or dying).

Confounding can be dealt with at the design stage of a study by *randomisation* (in a randomised controlled trial), *restriction*, or *matching* (in a case-control study). Alternatively, confounding variables can be controlled for at the analysis stage, by *stratification* (splitting the analysis e.g. by age group), *standardisation*, or *regression* (building a statistical model).

In Richiardi's case-control study, regression was used to control for the effect of smoking on lung cancer risk. The lung cancer risk associated with working as a docker or freight handler after controlling for the effect of smoking was reduced to 1.3 (95% confidence interval 0.9 to 1.9). Although the odds ratio is still higher (by 30%) for dockers or freight handlers, the confidence interval now spans 1 and so we can accept the null hypothesis that working as a docker or freight handler has no effect on lung cancer risk. This is because the higher odds reported for dockers or freight handlers could just have been found by chance. Smoking is therefore a confounder here, as it's associated with both the exposure (being a docker or freight handler) and the disease (lung cancer).

We can only adjust (control) for confounding factors if we have measured them. How often have you watched a TV news piece about an association between some potential risk factor and a disease and wondered, 'but could that be due to X instead?' Journalists rarely bother to talk about confounders.

Tutorial questions
The following questions will be undertaken in small groups, facilitated by a tutor. All the questions are designed to test your understanding of, and help you apply, the knowledge you will have learnt by reading the above worked examples, from listening to your tutor briefly explain the core concepts in the worked examples, and from the material covered in your lectures on the course so far. The questions should be worked though in groups; if you get stuck at any point please refer to the glossary at the end of this tutorial and ask your tutor for help.

Question 1 – Sampling distribution and confidence intervals
A study was conducted to assess whether hormone replacement therapy (HRT) conferred a protective effect on acute myocardial infarction *risk*. 1013 women with an acute myocardial infarction and 5000 women of a similar age range without acute myocardial infarction were asked whether or not they currently used HRT. 13.1% of the women who had had an MI used HRT, whereas 17.1% of the women who had not had an MI had used HRT. This study reported an *odds ratio* of 0.72 (95% *confidence interval* 0.59-0.88) for current or recent HRT use on acute myocardial infarction *risk* (Varas Lorenzo, 2000).

a) What type of study is this?

b) Why were 1013 women with an MI recruited instead of, say, 50? Why not 50,000?

c) Why were the 5000 "*controls*" (women without the *outcome* of interest, i.e. MI) chosen to have a similar age range as the "*cases*" (women with MI)?

d) What is the *null hypothesis* that this study is trying to disprove? Always be specific – don't just say "that there is no difference".

e) The 95% *confidence interval* for the *odds ratio* was 0.59-0.88. What does this mean?

f) For us to accept the *null hypothesis*, what would the 95% *confidence interval* look like? Give an example of its values.

g) What does the *odds ratio* of 0.72 mean in words, and how would you explain this *odds ratio* to someone taking HRT?

Question 2 – Dealing with confounding (in study design and analysis)
In a randomised controlled trial of patient self-monitoring of blood pressure in Birmingham general practices (McManus et al, 2005), 441 hypertensives were randomly allocated to either the usual monitoring by the practice (*control* group) or self-monitoring (intervention group). After six months, the intervention group reduced their systolic BP by an average of 4.3 mmHg (95% CI 0.8-7.9) more than the *control* group.

a) What is an appropriate distribution for a group of patients' BP?

b) What was the main *null hypothesis* for this study? Be specific, rather than just saying that "there is no difference".

c) Do we have evidence to reject the *null hypothesis*? What does this mean?

d) Why did the investigators randomly allocate patients to the two groups?

e) *Randomisation* was “stratified by diabetic status”. What does this mean and why was it done?

f) Other than diabetes, what other confounders might we want to control for?

g) The authors found that the intervention group had lost more weight and cut down their alcohol at the six-month follow-up stage. What is the relevance of this finding?

Question 3 – Understanding measures of association (and confounding)
To estimate the *incidence* of breast cancer in the UK population, records from the NHS breast screening programme (which screens women aged between 50 and 70) were explored. These data indicated that the *incidence* of breast cancer was 289 per 100,000 population.

a) What does "*incidence*" mean?

b) What can the *incidence* in this *sample* of the population tell us about the *incidence* in the whole UK female population?

c) A *null hypothesis* that the *incidence* of breast cancer in the UK female population aged 50-70 (289 per 100,000) is that same as the *incidence* in the UK female population aged 30-50 (90 per 100,000) gives a *p-value* of <0.0001. How would you interpret this *p-value*?

d) The *risk* of getting breast cancer if you are a woman aged 50-70 relative to the *risk* of getting breast cancer if you are a woman aged 30-50 is 3.20 (95% *confidence interval* 3.11-3.29). How would you interpret this *relative risk*?

e) The *odds* of being aged 50-70 if you have breast cancer compared with the *odds* of being aged 30-50 if you have breast cancer is also 3.2 (95% *confidence interval* 3.11-3.29). When the *odds*

f) The crude *relative risk* of breast cancer in women who are current users of HRT is 1.83 (95% CI 1.72-1.93), compared with the age-adjusted *relative risk* of 2.00 (1.91-2.09) (Beral, 2003). Which of these *risk* estimates would you consider to best reflect the *risk* of breast cancer associated with HRT use?

Question 4 – Relative risk vs attributable risk
An occupational study was carried out to investigate the effect of *exposure* to aromatic amines on bladder cancer *risk*. 6667 workers with potential *exposure* to aromatic amines were followed over 30 years to see what effect this *exposure* had on bladder cancer *risk*.

a) What type of study is this?

b) One quarter of the study population were exposed to aromatic amines, and the *risk* associated with this *exposure* on bladder cancer was found to be 296.94 (95% CI 41.45-2127.34). What does this *risk* measure tells us?

c) How would you explain this *risk* to someone with occupational *exposure* to aromatic amines?

d) The *population excess fraction* (excess fraction of bladder cancer due to aromatic amine *exposure* in the whole study population) is 98.7 percent. How would you interpret this figure?

e) One quarter of this study population are cigarette smokers. Cigarettes contain low doses of aromatic amines and have also been found to be associated with an excess *risk* of bladder cancer, with a *relative risk* of 5.11 (95% CI 3.42-7.64), and a *population excess fraction* of 50.7%.

f) Assume the same *risks* associated with occupational aromatic amine *exposure* (*relative risk* of ~297) and smoking (*relative risk* of ~5) in the occupational cohort apply to the whole population of England. In this England 'cohort', only 0.001% of the population has occupational exposure to aromatic amines, whilst 25% smoke. The *population excess fraction* is now 22.8% for aromatic amines, but remains at ~50% for smoking. Which *exposure* should be minimised to reduce *incidence* of bladder cancer in this population?

g) Assuming the *relative risk* of smoking on coronary heart disease mortality is ~2 (*population excess fraction* ~20%), and again taking the *relative risk* of smoking on bladder cancer in the population of England to be 5 (*population excess fraction* ~50%), and how is it that more deaths from coronary heart disease are attributed to smoking than bladder cancer *cases*?

h) What is the most useful measure of *risk* – the relative or the absolute (excess fraction) *risk*?

References

Beral, V. & Million Women, S. C. 2003, "Breast cancer and hormone-replacement therapy in the Million Women Study.[see comment][erratum appears in Lancet. 2003 Oct 4;362(9390):1160]", *Lancet*, vol. 362, no. 9382, pp. 419-427.

Malarcher, A. M., Schulman, J., Epstein, L. A., Thun, M. J., Mowery, P., Pierce, B., Escobedo, L., & Giovino, G. A. 2000, "Methodological issues in estimating smoking-attributable mortality in the United States", *American Journal of Epidemiology*, vol. 152, no. 6, pp. 573-584.

McManus RJ, Mant J, Roalfe A, Oakes RA, Bryan S, Pattison HM, Hobbs FDR. Targets and self monitoring in hypertension: randomised controlled trial and cost effectiveness analysis. Br Med J 2005 (Sep); 331: 493-498

Richiardi, L., Forastiere, F., Boffetta, P., Simonato, L., & Merletti, F. 2005, "Effect of different approaches to treatment of smoking as a potential confounder in a case-control study on occupational exposures", *Occupational & Environmental Medicine*, vol. 62, no. 2, pp. 101-104.

Varas-Lorenzo, C., Garcia-Rodriguez, L. A., Perez-Gutthann, S., & Duque-Oliart, A. 2000, "Hormone replacement therapy and incidence of acute myocardial infarction. A population-based nested case-control study", *Circulation*, vol. 101, no. 22, pp. 2572-2578.

Glossary

Words *in italics* are defined elsewhere in the glossary

Attributable risk – the *attributable risk* is a measure of ***exposure*** effect that indicates, on an absolute scale, how much greater the frequency of disease in the exposed group is compared with the unexposed, assuming the relationship between ***exposure*** and disease is causal (an important assumption). It is the difference between the *incidence rate* in the exposed and non exposed groups, i.e. it represents the *risk* attributable to the *exposure* of interest.

Attributable risk = Incidence in the exposed- Incidence in the unexposed

For example, if 20 out of 100 smokers got lung cancer (in a given period of time) compared with 5 out of 100 non-smokers, the *relative risk* (see below) would be 20/5 = 4, but the *attributable risk* would be (20 - 5)/100 = 15 per 100. This may also be expressed as an excess fraction; 15 per 100/20 per 100 = 75%. Of the 20 *cases* of lung cancer in the smoking population, 15 of them (75%) could be attributed to smoking. The *attributable risk* is especially useful in evaluating the impact of introduction or removal of risk factors. Its value indicates the number of *cases* of the disease among the exposed group that could be prevented if the *exposure* were completely eliminated.

Case – an individual with the *outcome* under study (in a *case-control study*). Epidemiological research is based on the ability to quantify the occurrence of disease in populations. This requires a clear definition of what is meant by a case. This could be a person who has the disease, health disorder, or suffers the event of interest (by "event" we mean a change in health status, e.g. death in studies of mortality or becoming pregnant in fertility studies). The epidemiological definition of a case is not necessarily the same as the clinical definition.

Case-control study – study in which individuals are selected on the basis of whether or not they have the *outcome* of interest; usually some relatively rare *outcome*. *Exposure* (risk factor) status is explored to establish whether the *exposure* is more common in the *case* (those that have the *outcome*) or *control* (those that do not have the *outcome*) group. This type of study always results in an *odds ratio*, for example comparing the *odds* of being exposed (e.g. a smoker) in those who had the *outcome* (e.g. pancreatic cancer), with the *odds* of being a smoker in those who did not have pancreatic cancer.

Cause – the key question in most medical research. Did *exposure* to electromagnetic radiation *cause* the leukaemia in children living near mobile phone masts? Did HRT *cause* the higher DVT *rates* in women taking it? Research works by trying to disprove alternative explanations (e.g. chance, *confounding*). If this can be done, then the relationship between the *exposure* and the *outcome* will be one of causation.

Count - the most basic measure of disease frequency is a simple count of affected individuals. The number (count) of *cases* that occurred in a particular population is of little use in comparing populations and groups. For instance, knowing that there were 100 *cases* of lung cancer in city A and 50 in city B does not tell us that lung cancer is more frequent in city A than B. There may simply be more people in city A. The number of *cases* may, however, be useful in planning services. For instance, if you wanted to set up an incontinence clinic, you would want to know the number of people with incontinence in your population.

Chi squared test – a statistical procedure for testing whether two proportions are similar (e.g. whether the proportion of lung cancer *cases* in males who smoke is significantly different to the proportion of lung cancer *cases* in males who do not smoke).

Cohort study – study in which individuals are selected on the basis of *exposure* status and are followed over a period of time to allow the frequency of occurrence of the *outcome* of interest in the exposed and non exposed groups to be compared. Take a group of people, note whether they've been exposed or not, observed them over time and wait for them to get ill, to die etc. This type of study typically produces a *relative risk.*

(95%) Confidence interval – an estimated range of values calculated from a given set of *sample* data which are likely to contain the 'true' population value. E.g. a range of values around a *relative risk* measure which would, in 95% of such studies, contain the 'true' *risk* (the true *risk* being the *relative risk* that would be obtained if the study had included the entire population of patients). By "contain (or 'span') the true value", we mean that the true value lies above the lower value of the *confidence interval* but below the upper values of the *confidence interval*. For example, for a *95% confidence interval* of 1.2 – 3.4, we can say that we are 95% confident that the true value of risk will not be lower than 1.2 and will not be higher than 3.4.
If we find that our confidence interval for the *relative risk* or *odds ratio* for group A compared with group B does not include 1, then we typically reject the null hypothesis of no difference. However, if our study is not on *rates* of disease or on proportions of patients *exposed* but is on a measure such as blood pressure or weight, we would typically reject the *null hypothesis* if the confidence interval for the average difference in blood pressure or weight between group A and group B does not include 0, not 1. Why is this? See entry for *null hypothesis*.

Confounding – a possible explanation for the study finding if *confounding variables* have not been taken into account in the study.

Confounding variable – a factor that is associated with both the *exposure* and *outcome* of interest. Common confounders include age, smoking, socio-economic deprivation. Smoking is a confounder because smoking tends to be more prevalent in people exposed to non-tobacco-related toxins and carcinogens, and also more prevalent in people with a range of diseases.

Control (as opposed to a *case*) – a person without the *outcome* under study (in a *case-control study*), or a person not receiving the intervention (in a clinical trial). The choice of an appropriate group of *controls* requires care, as we need to be able to draw useful comparisons between these *controls* and the *cases*/intervention group.

Exposure – when people have been 'exposed', they have been in contact with something that is hypothesised to have an effect on health e.g. tobacco, nuclear radiation, pesticides in food, HRT. Contact may be via any route: oral, inhalation, through the skin etc. These are typically called 'risk factors' of disease. We are interested in whether the *exposure* results in higher (or sometimes lower) *outcome rates*.

Incidence – the number of new *cases* of the *outcome* of interest occurring in a defined population over a define period of time. Note that this is not the same as *prevalence*, which includes new and old *cases*. *Incidence* measures events (a change from a healthy state to a diseased state).

$$\text{Incidence} = \frac{\text{Number of new cases of disease in a given time period}}{\text{Number of disease-free persons at the beginning of that time period}}$$

This measure of *incidence* can be interpreted as the probability, or risk, that an individual will develop the disease during a specific time period.

Matching - a method for "controlling for" (i.e. effectively removing) the effect of *confounding* at the design stage of a *case-control study*; *controls* are selected to have a similar distribution of potentially *confounding variables* to the *cases,* e.g. they are said to be "matched" for sex if there are similar proportions of men and women in both groups.

Normal distribution – a set of values and frequencies that describe many things in nature, at least approximately, e.g. height, weight, blood pressure. This symmetrical distribution (see Figure 1) is the basis of many statistical tests because, if you know the average value (usually called the mean) and the standard deviation, then you can draw every point of a *normal distribution* and you know what proportion of values are greater than (or less than) any given point, e.g. the % of men more than two metres tall. Some things are not normally distributed (e.g. proportions of anything, serum concentrations of electrolytes) but can be made to fit quite well after some simple mathematical trickery.

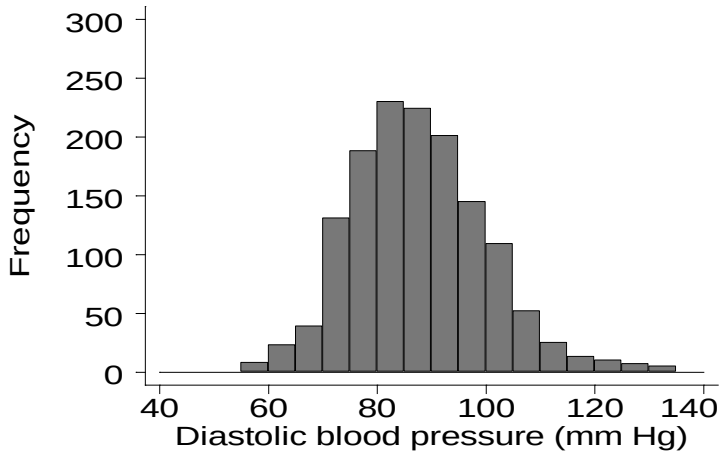


Figure 1. A *normal distribution* - from a study of diastolic blood pressure among men (*British Medical Journal*·1974; **3**: 600-3).

Null hypothesis – formulating a *null hypothesis* is the first stage in performing any statistical test. Typically, when two groups (A and B) are being compared, the *null hypothesis* that the statistical test tries to disprove is that there is no difference between the two groups in the measure being tested. If we are comparing rates, then the null hypothesis would be that *rate* A equals *rate* B, which means that the *relative risk* (*rate* A divided by *rate* B) equals 1. For *case-control studies*, the null hypothesis would be that the odds of *exposure* for group A equal the odds of *exposure* for group B, i.e. the *odds ratio* (*odds* of exposure for A divided by the *odds* of exposure for B) equals 1. A *statistical test* is then performed on the *relative risk* or the *odds ratio* and a *confidence interval* for it is derived. We can reject the null hypothesis if the *confidence interval* does not include the value expected under the null. In this case, the null has RR=1 or OR=1, so we would reject it if the *confidence interval* does not include 1.
However, for *normally distributed* variables such as blood pressure (BP) in Question 4, the null hypothesis would be that the average BP for group A equals the average BP

for group B, i.e. the difference between the two average BPs equals 0. The *statistical test* would then be performed on this difference in average BPs and the resulting *confidence interval* would also relate to the difference in average BPs. We therefore would reject the null hypothesis if the *confidence interval* did not include 0, which is the value expected under the null.
If, when faced with a *confidence interval* around some measure and wondering whether to reject the null hypothesis or not, you can't remember whether it should include 1 or 0, always think in terms of what value the null hypothesis expects your measure to have and then see if that value falls within the range of values covered by the *confidence interval*.

Odds – the *odds* is another way to express probability, e.g. the *odds* of *exposure* is the number of people who have been exposed divided by the number of people who have not been exposed. The mathematical relationship between odds and probability is:
Odds = probability / (1 – probability)

Odds ratio – the *relative risk* can be calculated from *cohort studies*, since the *incidence* of disease in the exposed and non-exposed is known. In *case-control studies*, however, the subjects are selected on the basis of their disease status (sample of subjects with a particular disease (*cases*) and sample of subjects without that disease (*controls*)), not on the basis of exposure. Therefore, it is not possible to calculate the incidence of disease in the exposed and non-exposed individuals. It is, however, possible to calculate the *odds of exposure*. The *odds ratio* (of exposure) is the ratio between two odds, e.g. the *odds* of *exposure* in the *case* s divided by the *odds* of *exposure* in the *control*s.

$$\text{Odds ratio} = \frac{\text{Odds of exposure in the diseased group (cases)}}{\text{Odds of exposure in the disease-free group (controls)}}$$

This ratio is the measure reported in *case-control studies* instead of the *relative risk*. It can be mathematically shown that the *odds ratio* of exposure is generally a good estimate of the *relative risk*. An *odds ratio* of 1 tells us that *exposure* is no more likely in the *cases* than *controls* (which implies that *exposure* has no effect on *case*/*control* status); an *odds ratio* greater than 1 tells us that *exposure* is more likely in the *case* group (which implies that *exposure* might increase the *risk* of the disease). An *odds ratio* less than 1 tells us that *exposure* is less likely in the *case* group (which implies that *exposure* might have a protective effect).

Outcome – the event or main quantity of interest in a particular study, e.g. death, contracting a disease, blood pressure.

Population attributable risk (also known as the population excess risk) – a measure of the *risk* of *outcome* in the study population which is attributable to the *exposure* of interest.

Population excess fraction (also known as the population attributable fraction) – a measure of the proportion (fraction) of the *cases* observed in the study population attributable to the *exposure* of interest.

Prevalence – the number of *cases* of an *outcome* of interest in a defined population at a particular point of time, hence it is often called *point prevalence.* This includes both new (also called "incident") *cases* and existing *cases*.

$$\text{Point prevalence} = \frac{\text{Number of cases in a defined population at one point in time}}{\text{Number of persons in a defined population at the same point in time}}$$

p-value – the probability of obtaining the study result (*relative risk*, *odds ratio* etc) if the *null hypothesis* is true. The smaller the p-value, the easier it is for us to reject the *null hypothesis* and accept that the result was not just due to chance. A *p-value* of <0.05 means that there is only a very small chance of obtaining the study result if the *null hypothesis* is true, and so we would usually reject the null. Such as result is commonly called "statistically significant". A *p-value* of >0.05 is usually seen as providing insufficient evidence against the *null hypothesis*, so we accept the null.

Randomisation – a method for ensuring that both groups in a clinical trial (i.e. those receiving the intervention and those not receiving the intervention (*controls*)), have similar proportions of *confounding variables*, such as age.

Rate and risk – these words are often taken to mean the same thing (though to some epidemiological purists they are not always the same). We talk of someone's *risk*/chance/probability of getting a disease (or getting pregnant or dying etc.) and a population having a disease *rate*. Both terms imply a proportion, i.e. the number of people with the *outcome* of interest divided by the total number of people at *risk* of the *outcome*.

Regression - a method for controlling the effect of *confounding* at the analysis stage of a study - statistical modelling is used to control for one or many *confounding variables*.

Relative risk – the *relative risk* is used as a measure of association between an exposure and disease. It is the ratio of the *incidence rate* in the *exposed* group and the *incidence rate* in the non-exposed group.

$$\text{Relative risk} = \frac{\text{Incidence in the exposed group}}{\text{Incidence in the unexposed group}}$$

For example, the proportion of people with high cholesterol who developed ischaemic heart disease divided by the proportion of people with normal cholesterol who developed ischaemic heart disease. A value of 1.0 indicates that the incidence of disease in the exposed and the unexposed are identical and thus the data shows no association between the exposure and the disease. A value greater than 1.0 indicates a positive association or an increased risk among those exposed to a factor. Similarly, a relative risk less than 1.0 means there is an inverse association or a decreased risk among those exposed, i.e. the exposure is protective.

Restriction – a method for controlling the effect of *confounding* at the design stage of a study, e.g. by including patients in a clinical trial only between the ages of 18 and 65 without pre-existing illness so that the results of the trial are not confused ('confounded') by different levels of age or morbidity in the two treatment groups.

Sample – a relatively small number of observations (or patients) from which we try to describe the whole population from which the *sample* has been taken. Typically, we calculate the mean for the *sample* and use the *confidence interval* to describe the range within which we think the population mean lies. This is one of the absolutely key concepts behind all medical research (and much non-medical research too).

Standardisation - a method for controlling the effect of *confounding* at the analysis stage of a study. Used to produce a Standardised Mortality Ratio, a commonly used measure in epidemiology.

Statistical test – the only way to decide whether the results of your analysis, e.g. your measure for group A compared with your measure for group B, are likely to be due to chance or could be real. The procedure for doing a statistical test is to take one value

representing the observed difference in your study between groups A and B and compare that value against tables of an appropriate mathematical distribution such as the *normal distribution* to see how extreme it is (we use computers instead of printed tables, thankfully, these days). For example, to see if someone is unusually tall, we would need to compare their height with a normal distribution with the mean and standard distribution taken from members of the population of the same age and sex. This would be done by subtracting the population mean from the person's height and dividing by the population standard deviation and looking up the result (called the "test statistic") in a table of the standard normal distribution (so-called because it has a mean of 0 and standard deviation of 1) to find out what proportion of values are greater than this. This proportion is therefore the proportion of the population who are taller than the person. Something similar is routinely done on infants to monitor their growth.

Stratification - a method for controlling the effect of *confounding* at the analysis stage of a study - *risks* are calculated separately for each category of *confounding variable*, e.g. each age group and each sex separately.

Tutorial 2

Critical appraisal of medical evidence

Dr Mireille B. Toledano and Dr Claire Robertson

Learning outcomes

- To learn how to read a paper in a scientific journal
- To be able to review and critically appraise medical evidence
- To be able to present critical appraisal findings to lecturers and peers

This tutorial is designed to help you understand how to read and interpret the evidence presented in papers published in medical journals.

Based on feedback from last years' students, we have changed the way we teach critical appraisal of medical evidence. Students felt that they would benefit from small group tutorial sessions where a tutor would be available to help clarify important points, answer questions, and facilitate discussion.

Why are critical appraisal skills important in medicine?

- They are a core part of clinical practice – finding, making sense of, and applying new research evidence to enhance practice
- They help to improve communication with patients – providing informed advice to patients asking about new treatments they have seen on the internet increasing survival or not etc
- They are a key skill that students are expected to use throughout the MBBS – e.g. in your Problem Based-Learning (PBL) groups, your BSc year, on ward rounds etc

How does this tutorial fit into the rest of the EIP course?

- This tutorial is both an opportunity for you to practice your critical appraisal skills and also *consolidate* and *apply* all the material you will have learnt in prior sessions on this course (i.e. core concepts of evidence-based medicine, study designs, interpretation of statistical findings etc) because all this knowledge is a necessary foundation to conduct an informed critical appraisal.
- The tutorial is directly linked to Lecture 13 (Introduction to critical appraisal of medical evidence) where you were taught core critical appraisal skills by Dr Claire Robertson.
- *All material taught in this tutorial will be included in your examinations.*

The tutorial is focused upon appraising the following two published papers investigating the effects of vitamin supplements on cardiovascular disease and mortality. Both of these papers are provided for you in this handbook (following this tutorial):

Paper 1: Pocobelli G, Peters U, Kristal AR, White E. Use of supplements of multivitamins, vitamin C, and vitamin E in relation to mortality. *Am J Epidemiol 2009:170; 472-483.*

Paper 2: Lee IM, Cook NR, Gaziano JM, Gordon D, Ridker PM, Manson JE, Hennekens CH, Buring JE. Vitamin E in the primary prevention of cardiovascular

disease and cancer: the Women's Health Study: a randomized controlled trial. *JAMA 2005: 294(1); 56-65*

You should read these two papers before the tutorial, so that you have sufficient time to work through the critical appraisal checklist (below) in your groups during the timetabled session with your tutor.

Format of the tutorial

- Each tutorial group will be asked to split into two sub-groups
- Each sub-group will be asked to read through and critically evaluate **one** of these papers. You should use as a guide:
    - **The general critical appraisal checklist** (which follows the format you will have been taught in Lecture 13) that you can use to evaluate ANY type of study design (see below).
    - **AND, for paper 1**, you should also consider the STROBE evaluation checklist specific to observational (cohort) studies (which will have been mentioned to you in Lecture 13), provided in this handbook following this tutorial.
    - **AND, for paper 2**, you should also consider the CONSORT evaluation checklist specific to clinical trials (which will have been mentioned to you in Lecture 13), provided in this handbook following this tutorial.
- The full group will then reconvene and each sub-group will present a critical appraisal of the findings of their paper to their peers and tutor.
- A tutor-led discussion on how to interpret the evidence in these papers in a wider context will then follow.

Critical appraisal checklists

1. General critical appraisal checklist (*further details on each of these checklist points can be found in the summary of Lecture 13)*

Summarise the paper first: with a sentence for each of the following:

- Why did they do it?
- What did they do?
- What did they find?
- What did they conclude?

Then consider the following:

Question
Design
Population
Methods
Analysis
Confounding
Bias
Ethics
Interpretation

The specific checklists for each type of study design are provided at the back of this course guide. For further information on these, you may also wish to look at:

STARD: further information at: http://www.stard-statement.org/

STROBE: further information at: http://www.strobe-statement.org/

MOOSE: further information at: http://jama.ama-assn.org/cgi/content/full/283/15/2008

CONSORT: further information at: http://www.consort-statement.org/

QUORUM-PRISMA: further information at: http://www.equator-network.org/resource-centre/library-of-health-research-reporting/reporting-guidelines/systematic-reviews-and-meta-analysis/

Recommended reading:

Ward H, Toledano M.B, Shaddick G, Davies B, Elliott P*, Oxford Handbook of* Epidemiology *for Clinicians (2012),* Oxford University Press

- Chapter 4, pages 78, 82-84, 86

Trisha Greenhalgh (2001). *How to read a paper: The basics of evidence based medicine.* BMJ.

American Journal of Epidemiology
Vol. 170, No. 4
DOI: 10.1093/aje/kwp167
Advance Access publication July 13, 2009

Original Contribution

Use of Supplements of Multivitamins, Vitamin C, and Vitamin E in Relation to Mortality

Gaia Pocobelli, Ulrike Peters, Alan R. Kristal, and Emily White

Initially submitted March 6, 2009; accepted for publication May 22, 2009.

In this cohort study, the authors evaluated how supplemental use of multivitamins, vitamin C, and vitamin E over a 10-year period was related to 5-year total mortality, cancer mortality, and cardiovascular disease (CVD) mortality. Participants ($n = 77{,}719$) were Washington State residents aged 50–76 years who completed a mailed self-administered questionnaire in 2000–2002. Adjusted hazard ratios and 95% confidence intervals were computed using Cox regression. Multivitamin use was not related to total mortality. However, vitamin C and vitamin E use were associated with small decreases in risk. In cause-specific analyses, use of multivitamins and use of vitamin E were associated with decreased risks of CVD mortality. The hazard ratio comparing persons who had a 10-year average frequency of multivitamin use of 6–7 days per week with nonusers was 0.84 (95% confidence interval: 0.70, 0.99); and the hazard ratio comparing persons who had a 10-year average daily dose of vitamin E greater than 215 mg with nonusers was 0.72 (95% confidence interval: 0.59, 0.88). In contrast, vitamin C use was not associated with CVD mortality. Multivitamin and vitamin E use were not associated with cancer mortality. Some of the associations we observed were small and may have been due to unmeasured healthy behaviors that were more common in supplement users.

ascrobic acid; cohort studies; coronary disease; dietary supplements; mortality; neoplasms; vitamin E; vitamins

Abbreviations: CI, confidence interval; CVD, cardiovascular disease; ICD-10, *International Classification of Diseases*, Tenth Revision; PHS, Physicians' Health Study; RR, relative risk; WACS, Women's Antioxidant Cardiovascular Study.

Free radicals are present in human cells both as a normal consequence of energy metabolism (1, 2) and as a consequence of exposure to exogenous factors such as cigarette smoke (1, 2). Laboratory studies have documented damage by free radicals, known as oxidative damage, to DNA (3, 4), proteins (5), and lipids (1). Because this type of damage is also associated with disease—for example, DNA damage (3, 4) and the occurrence of cancer (6) and lipid peroxidation (1) and the development of atherosclerosis (7)—attention has focused on the respective roles of free radicals and antioxidants in disease causation and prevention. Antioxidants, such as vitamins C and E, may be capable of preventing oxidative damage in human cells because they are strong electron donors and therefore are relatively quick to react with a free radical (1, 2).

Multivitamin and vitamin C and E supplements are commonly used in the United States (8). Whether or not use of these supplements is related to mortality is an important consideration in an evaluation of whether to initiate or continue their use. Currently, there is no clear evidence that taking multivitamins or vitamin C or E supplements delays mortality or, more specifically, reduces a person's risk of death from cardiovascular disease (CVD) or cancer. Findings from cohort studies of these associations are inconsistent (9–17), and findings from meta-analysis of randomized trials tend to show no benefit (18–22), although there are no published results from randomized trials of common multivitamin formulations and risk of death.

Randomized trials have the advantage of protecting against confounding by unmeasured variables, but their ability to detect an association may be limited by incomplete

Correspondence to Gaia Pocobelli, Fred Hutchinson Cancer Research Center, M3-B232, 1100 Fairview Avenue North, Seattle, WA 98109-1024 (e-mail: gpocobel@u.washington.edu).

adherence in the study arms formed by randomization (23), a supplement dose that is not in the range needed (23), or a duration of use that is too short (23) to affect a person's risk of death.

In this cohort study, which was specifically designed to recruit supplement users and to measure their use of supplements, we evaluated the association between intake of multivitamins and vitamin C and E supplements in the 10 years before baseline and risk of total mortality, CVD mortality, and cancer mortality during the 5 years after baseline.

MATERIALS AND METHODS

Study population

The Vitamins and Lifestyle Study is a prospective study of men and women aged 50–76 years in western Washington State. The proposal for this study was approved by the institutional review board of the Fred Hutchinson Cancer Research Center (Seattle, Washington). The study's design was previously described in detail (24). Briefly, 364,418 persons identified from a commercial mailing list were mailed a cover letter that targeted supplement users and a 24-page sex-specific baseline questionnaire. Included in the cover letter was a statement that this was a study of how "vitamin supplements, certain foods, and physical activity can influence your risk of cancer," and in pilot testing, inclusion of this statement led to an increased frequency of participation by supplement users. Between October 2000 and December 2002, a total of 77,719 persons returned a questionnaire that passed eligibility and quality control checks. For the present analysis, we excluded 1 participant with no follow-up time and 45 participants who reported having a malabsorption condition (e.g., a prior gastroplasty) at baseline (these conditions are associated with decreased nutrient absorption); this left 77,673 participants.

Ascertainment of supplement use and potential confounders

Supplement use. For each type of supplement used, information was obtained on the duration, frequency, and dose per day on the days the supplement was taken. Ever use of a supplement was defined as use at least once per week, for a year, during the 10-year period before baseline.

A multivitamin was defined as a mixture containing at least 10 vitamins and/or minerals. Information was obtained on the brand of multivitamin currently used and the brand most commonly used in the past. Ten-year average frequency of multivitamin use (days/week) was computed as "duration (years)/10 (years) × frequency (days/week)."

We also computed 10-year average dose per day of vitamins C and E from single supplements (including mixtures other than multivitamins) plus multivitamins. To do so, we estimated the amounts of vitamins C and E in each subject's brand of multivitamin based on the *Physicians' Desk Reference for Nonprescription Drugs and Dietary Supplements 2002* (25) or the amount reported by the manufacturer or participant. Ten-year average doses of supplemental vitamin C and vitamin E (mg/day) were then computed as "duration (years)/10 (years) × frequency (days/week)/7 (days/week) × dose per day (mg/day)," summed over individual supplements and multivitamins.

Potential confounders. The following characteristics were considered a priori to be potential confounders because they might have been associated with supplement use and mortality: sex, age, race/ethnicity, marital status, education, recency/dose of smoking, alcohol intake, average physical activity in the 10 years before baseline (26), body mass index (weight (kg)/height $(m)^2$), age at menopause, estrogen therapy, estrogen plus progestin therapy, use of regular or extra-strength aspirin in the previous 10 years, use of other nonaspirin nonsteroidal antiinflammatory medication in the previous 10 years, current use of cholesterol-lowering medication, receipt of a prostate-specific antigen test in the previous 2 years, receipt of a mammogram in the previous 2 years, receipt of a sigmoidoscopy in the previous 10 years, self-rated health, health history (see below), mother's and father's ages at death, and diet (see below). For body mass index and alcohol intake, we adjusted for measures at 45 years of age rather than at baseline because the former were more strongly related to mortality.

Diet in the year before baseline was measured with a modified version of the food frequency questionnaire used in the Women's Health Initiative (27). Based on the components of diet recommended by the US Dietary Guidelines Advisory Committee (28), selected dietary variables were evaluated for their relation to mortality. The following variables were related to mortality and were included in the final statistical models: percentage of energy derived from *trans* fat, percentage of energy derived from saturated fat, daily number of servings of fruits, and daily number of servings of vegetables (excluding potatoes).

To adjust for health history at baseline, we created a morbidity score. Sex-specific age-adjusted Cox proportional hazards models (29) were used to determine the hazard ratio for death for each of 23 conditions for men and each of 27 conditions for women, modeled simultaneously (see footnote "c" in Table 1 for a list of the conditions). Using the coefficients from these models, we assigned each subject a morbidity score that was the natural logarithm of the hazard ratio for death based on his/her particular set of comorbid conditions as compared with persons with no comorbid conditions.

Ascertainment of death

We linked the cohort to the Washington State Death Certificate System to identify deaths occurring through December 31, 2006 (n = 3,535) (24). Additional deaths were identified from the Social Security Death Index (n = 37), linkage with the western Washington Surveillance, Epidemiology, and End Results cancer registry (n = 2), and notification by relatives (n = 3), for a total of 3,577 deaths (24).

The date of death was available for all deaths. Information on cause of death was available only for deaths identified through the Washington State Death Certificate System. It was determined from the underlying cause of death coded using the *International Classification of*

Table 1. Total Mortality Rates and Hazard Ratios for Total Mortality According to Participant Characteristics at Baseline, Vitamins and Lifestyle Study, Western Washington State, 2000–2006

Characteristic	No. of Subjects (n = 77,673)	%	Person-Years of Follow-up (n = 387,801)[a]	%	No. of Deaths (n = 3,577)	%	Mortality Rate[b]	Sex- and Age-Adjusted Hazard Ratio	95% Confidence Interval
Sex									
Female	40,308	52	202,169	52	1,514	42	7.49	1.00	Referent
Male	37,365	48	185,633	48	2,063	58	11.11	1.50	1.40, 1.60
Age at baseline, years									
50–54	17,952	23	91,245	24	263	7	2.88	1.00	Referent
55–59	17,566	23	87,978	23	419	12	4.76	1.65	1.42, 1.93
60–64	14,121	18	70,450	18	533	15	7.57	2.61	2.25, 3.02
65–69	12,834	17	63,647	16	789	22	12.40	4.26	3.71, 4.90
70–76	15,200	20	74,481	19	1,573	44	21.12	7.37	6.47, 8.40
Race/ethnicity									
White	71,096	92	355,127	92	3,276	92	9.22	1.00	Referent
Hispanic	669	1	3,330	1	16	0	4.80	0.70	0.42, 1.11
Black	990	1	4,872	1	61	2	12.52	1.39	1.08, 1.79
American Indian/Alaska Native	1,152	1	5,729	1	59	2	10.30	1.28	0.99, 1.65
Asian or Pacific Islander	1,937	2	9,751	3	66	2	6.77	0.78	0.61, 0.99
Other/missing data	1,829	2	8,992	2	99	3	11.01	1.06	0.86, 1.29
Marital status									
Married	57,212	74	286,458	74	2,390	67	8.34	1.00	Referent
Living with a partner	1,986	3	10,010	3	76	2	7.59	1.31	1.04, 1.64
Separated or divorced	8,943	12	12,521	11	442	12	9.99	1.54	1.39, 1.72
Widowed	5,570	7	44,250	7	469	13	17.07	1.46	1.32, 1.63
Never married	2,514	3	27,470	3	119	3	9.50	1.48	1.23, 1.78
Missing data	1,448	2	7,092	2	81	2			
Education	76,225								
Grade school/some high school	2,702	4	13,194	3	295	8	22.36	1.00	Referent
High school graduation/General Equivalency Diploma	12,747	16	63,471	16	825	23	13.00	0.75	0.66, 0.86
Some college/technical school	29,237	38	145,763	38	1,388	39	9.52	0.66	0.58, 0.75
College graduation	18,677	24	93,655	24	656	18	7.00	0.48	0.41, 0.55
Advanced degree	12,978	17	65,205	17	334	9	5.12	0.36	0.31, 0.42
Missing data	1,332	2	6,513	2	79	2			
Morbidity score[c]									
Level 1 ($\leq$0)	35,466	46	179,929	47	616	17	3.42	1.00	Referent
Level 2 ($>$0–$<$0.5)	27,916	36	139,999	36	1,015	29	7.25	1.70	1.54, 1.88
Level 3 (0.5–$<$1.0)	7,733	10	37,899	10	644	18	16.99	3.62	3.24, 4.05
Level 4 (1.0–$<$1.5)	3,978	5	18,827	5	586	16	31.13	6.13	5.46, 6.89
Level 5 (1.5–$<$2.0)	1,397	2	6,203	2	334	9	53.84	10.20	8.91, 11.69
Level 5 (2.0–$<$2.5)	503	1	2,116	1	157	4	74.20	14.09	11.80, 16.82
Level 6 (2.5–$<$3.0)	256	0	960	0	117	3	121.88	22.62	18.52, 27.62
Level 7 ($\geq$3.0)	192	0	715	0	89	3	124.48	23.41	18.71, 29.29
Missing data	232	0	1,153	0	19	1			

[a] Because of rounding, numbers of person-years for each variable do not always sum to exactly 387,801.

[b] Number of deaths per 1,000 person-years.

[c] The following conditions, categorized as yes or no, were modeled simultaneously in sex-specific and age-adjusted models to obtain the morbidity score: current use of medication for depression or anxiety; current use of blood pressure medication; a history of lung cancer, colon cancer, bladder cancer, leukemia, pancreatic cancer, non-Hodgkin's lymphoma, melanoma, prostate cancer, breast cancer, cervical cancer, uterine cancer, ovarian cancer, or all other cancers combined; coronary heart disease (defined as a previous heart attack, coronary bypass surgery, angioplasty, or diagnosis of angina); stroke; congestive heart disease; rheumatoid arthritis; diabetes; viral hepatitis; cirrhosis of the liver; other chronic liver disease; emphysema; chronic bronchitis or chronic obstructive pulmonary disease; kidney disease; ulcerative colitis or Crohn's disease; Parkinson's disease; and osteoporosis in women.

Diseases, Tenth Revision (ICD-10) (30). We classified deaths as being due to CVD (ICD-10 codes I00–I15, I20–I52, and I60–I99), cancer (ICD-10 codes C00–D48), or other causes.

Statistical analysis

Cox proportional hazards regression (29), with age as the time variable, was used to determine the hazard ratio for death (and 95% confidence interval) associated with supplement use, with adjustment for potential confounders. Participants were considered to be at risk for mortality from their age at completion of the baseline questionnaire through their age at death ($n = 3{,}577$) or age at censoring (withdrawal from the study ($n = 22$), moving out of Washington State ($n = 3{,}224$), or December 31, 2006 ($n = 70{,}850$)). We identified participants who had moved through linkage to the National Change of Address file, with follow-up by mail or phone (24).

To reduce the numbers of participants dropped from analyses because of missing data, we included a ''missing'' category for most confounders; nonetheless, 7%–12% of participants were excluded from each analysis because of missing data on exposure or confounding factors.

The statistical significance of the supplement variable was tested using a likelihood ratio test for trend with the exposure variable categorized in ordinal form. Because this test assumes a log-linear relation between the hazard ratio for mortality and the supplement use variable, we first tested for nonlinearity in this relation. To do so, we compared the model with the supplement variable categorized as a dummy variable with the model with the supplement variable categorized as an ordinal variable, and if they differed at a P value of 0.05, the test for trend was not conducted.

Statistical tests of interaction were performed using a likelihood ratio test comparing models with and without the interaction terms. The interaction terms were the products of the supplement use variable, coded as an ordinal variable, and the modifier variable, coded as a dummy variable.

We also determined the hazard ratios for death from CVD, cancer, and all other causes combined associated with supplement use. Analyses of death from CVD were stratified by history of CVD, and results were adjusted for potential confounders (see table footnotes). Analyses of death from cancer were stratified by history of cancer (excluding nonmelanoma skin cancer), and results were adjusted for potential confounders (see table footnotes).

RESULTS

During 387,801 person-years of follow-up, 3,577 deaths occurred among 77,673 participants (9.22 deaths per 1,000 person-years) (Table 1). Sixty-six percent of participants had ever used multivitamins, 47% had used a vitamin C supplement, and 48% had used a vitamin E supplement (Table 2). After multivariate adjustment, multivitamin use was not associated with risk of total mortality, whether evaluated by duration, frequency during period of use, or 10-year average frequency of use. Vitamin C use was associated with a small decreased risk of total mortality when evaluated by duration of use (P-trend = 0.019), average dose on days taken (P-trend = 0.023), and 10-year average daily dose (P-trend = 0.032). The hazard ratio comparing persons in the third tertile ($\geq$322.1 mg/day) of 10-year average daily dose with nonusers was 0.89 (95% confidence interval (CI): 0.81, 0.98). Vitamin E use was also associated with a small decreased risk of total mortality when it was evaluated by average dose on days taken (P-trend = 0.010) and 10-year average daily dose (P-trend = 0.008). The hazard ratio comparing persons in the third tertile ($\geq$215.1 mg/day) of 10-year average daily dose with nonusers was 0.89 (95% CI: 0.81, 0.98).

We also evaluated whether the hazard ratios for total mortality associated with 10-year average daily dose of vitamins C and E varied according to several participant characteristics (Table 3). Among never smokers, risk of total mortality was inversely related to use of supplemental vitamin C (hazard ratio = 0.76, 95% CI: 0.63, 0.92) and vitamin E (hazard ratio = 0.80, 95% CI: 0.66, 0.97) when comparing the highest tertile of use with nonuse, whereas there were no associations among current/recent smokers. Risk of total mortality was also inversely related to use of vitamins C and E among persons with a body mass index of 30 or greater; the respective hazard ratios were 0.76 (95% CI: 0.57, 1.01) and 0.78 (95% CI: 0.58, 1.04) when comparing the highest tertile of use with nonuse, whereas there were no associations among persons with a body mass index less than 25. Additionally, risk of total mortality was inversely related to use of vitamins C and E among persons who consumed less than the median daily number of servings of fruits and vegetables but not in persons who consumed at least the median number of servings per day. When results were stratified by age (data not shown), sex, alcohol use at age 45 years (data not shown), or morbidity score, the hazard ratios associated with increasing dose for both vitamin C and vitamin E did not vary markedly.

We also evaluated risk of death from CVD, cancer, and all other causes combined in relation to 10-year average daily dose of multivitamins, vitamin C, and vitamin E (Table 4). Multivitamin use was inversely associated with risk of CVD mortality (P-trend = 0.019) but not mortality from cancer or from all other causes combined. Overall, vitamin C use was not associated with CVD mortality, but it was inversely associated with risk among persons with a history of CVD at baseline (P-trend = 0.036). It was also associated with cancer mortality among persons in the third tertile of use ($\geq$322.1 mg/day) as compared with nonusers; however, there was no evidence of a dose-response relation. Vitamin E use was inversely related to risk of CVD mortality (P-trend = 0.001) only.

DISCUSSION

Our results should be interpreted in the context of several limitations. Although we adjusted for many factors associated with both supplement use and mortality, confounding by unmeasured factors may have occurred. For example, supplement users may be more likely than nonusers to participate in screening or comply with treatment for disease.

Table 2. Total Mortality Rates and Hazard Ratios for Total Mortality Associated With Supplement Use During the 10 Years Before Baseline, Vitamins and Lifestyle Study, Western Washington State, 2000–2006

Supplement	Subjects (n = 77,673)		Person-Years of Follow-up (n = 387,801)[a]		Deaths (n = 3,577)		Mortality Rate[b]	Sex- and Age-Adjusted HR	95% CI	Multivariate-Adjusted HR[c]	95% CI
	No.	%	No.	%	No.	%					
Multivitamins											
Duration of use, years											
None	25,759	33	128,786	33	1,266	35	9.83	1.00	Referent	1.00	Referent
1–3	9,009	12	45,148	12	388	11	8.59	1.02	0.91, 1.15	1.02	0.90, 1.14
4–6	8,931	11	44,619	12	372	10	8.34	0.92	0.82, 1.03	0.93	0.85, 1.09
7–9	6,337	8	31,522	8	306	9	9.71	1.00	0.88, 1.13	1.09	0.96, 1.24
≥10	24,471	32	121,960	31	1,091	31	8.95	0.83	0.77, 0.90	0.97	0.89, 1.06
Missing data	3,166	4	15,766	4	154	4	9.77				
P-trend								0.001		0.644	
Frequency of use during period of use, days/week											
None	25,759	33	128,786	33	1,266	35	9.83	1.00	Referent	1.00	Referent
1–2	1,675	2	8,398	2	74	2	8.81	0.98	0.77, 1.23	0.93	0.72, 1.20
3–4	3,190	4	15,999	4	102	3	6.38	0.80	0.65, 0.98	0.90	0.73, 1.12
5–6	7,911	10	39,792	10	214	6	5.38	0.66	0.57, 0.76	0.86	0.74, 1.00
7	34,000	4	169,285	44	1,650	46	9.75	0.93	0.86, 1.00	1.02	0.95, 1.11
Missing data	5,138	7	25,540	7	271	8	10.61				
P-trend								N/A[d]		0.690	
Ten-year average frequency of use, days/week											
None	25,759	33	128,786	33	1,266	35	9.83	1.00	Referent	1.00	Referent
>0–2	12,405	16	62,063	16	551	15	8.88	1.02	0.93, 1.13	1.02	0.92, 1.14
3–5	10,541	14	52,707	14	411	11	7.80	0.85	0.76, 0.95	0.94	0.83, 1.05
6–7	26,845	35	133,639	34	1,254	35	9.38	0.87	0.81, 0.94	1.00	0.92, 1.09
Missing data	2,123	3	10,606	3	95	3	8.96				
P-trend								0.001		0.873	
Vitamin C											
Duration of use[e], years											
None	41,490	53	206,723	53	2,063	58	9.98	1.00	Referent	1.00	Referent
1–3	6,906	9	34,617	9	294	8	8.49	0.95	0.84, 1.08	1.02	0.90, 1.16
4–6	6,344	8	31,815	8	262	7	8.24	0.89	0.78, 1.01	0.97	0.85, 1.11
7–9	4,296	6	21,540	6	165	5	7.66	0.77	0.65, 0.90	0.86	0.72, 1.01
≥10	15,366	20	76,722	20	647	18	8.43	0.77	0.70, 0.84	0.91	0.83, 1.00
Missing data	3,271	4	16,385	4	146	4	8.91				
P-trend								<0.001		0.019	
Dose[e] on days taken, mg/day											
None	41,490	53	206,723	53	2,063	58	9.98	1.00	Referent	1.00	Referent
60–250	4,385	6	21,848	6	209	6	9.57	0.93	0.80, 1.07	1.02	0.88, 1.19
500	14,850	19	74,418	19	590	16	7.93	0.75	0.69, 0.83	0.90	0.81, 0.99
1,000	11,768	15	59,080	15	448	13	7.58	0.80	0.72, 0.88	0.92	0.82, 1.02
1,500	2,484	3	12,397	3	104	3	8.39	0.91	0.75, 1.11	0.92	0.75, 1.13
Missing data	2,696	3	13,335	3	163	5	12.22				
P-trend								<0.001		0.023	

Table continues

Table 2. Continued

Supplement	Subjects (n = 77,673)		Person-Years of Follow-up (n = 387,801)[a]		Deaths (n = 3,577)		Mortality Rate[b]	Sex- and Age-Adjusted HR	95% CI	Multivariate-Adjusted HR[c]	95% CI
	No.	%	No.	%	No.	%					
Ten-year average dose[f], mg/day											
None	20,713	27	103,444	27	1,063	30	10.28	1.00	Referent	1.00	Referent
Tertile 1 (2.6–60.0)	19,334	25	96,195	25	925	26	9.62	0.92	0.84, 1.00	0.97	0.89, 1.07
Tertile 2 (60.1–322.0)	18,283	24	91,439	24	784	22	8.57	0.82	0.75, 0.90	0.97	0.88, 1.07
Tertile 3 (322.1–1,750.0)	18,710	24	93,613	24	762	21	8.14	0.73	0.66, 0.80	0.89	0.81, 0.98
Missing data	633	1	3,110	1	43	1	13.83				
P-trend								<0.001		0.032	
Vitamin E											
Duration of use[e], years											
None	40,445	52	201,496	52	2,030	57	10.07	1.00	Referent	1.00	Referent
1–3	9,680	12	48,914	13	354	10	7.24	0.77	0.69, 0.87	0.89	0.79, 1.00
4–6	8,490	11	42,505	11	322	9	7.58	0.74	0.65, 0.83	0.83	0.73, 0.94
7–9	4,704	6	23,470	6	205	6	8.73	0.80	0.69, 0.92	1.00	0.86, 1.16
≥10	11,501	15	57,137	15	530	15	9.28	0.74	0.67, 0.82	0.89	0.80, 0.99
Missing data	2,853	4	14,280	4	136	4	9.52				
P-trend								N/A		N/A	
Dose[e] on days taken, mg/day											
None	40,445	52	201,496	52	2,030	57	10.07	1.00	Referent	1.00	Referent
30–200	3,664	5	18,392	5	150	4	8.16	0.81	0.68, 0.95	0.85	0.71, 1.01
400	23,267	30	116,626	30	926	26	7.94	0.70	0.65, 0.76	0.88	0.81, 0.96
600–800	7,079	9	35,413	9	299	8	8.44	0.84	0.74, 0.95	0.91	0.80, 1.04
Missing data	3,218	4	15,874	4	172	5	10.84				
P-trend								N/A		0.010	
Ten-year average dose[f], mg/day											
None	20,259	26	101,130	26	1,050	29	10.38	1.00	Referent	1.00	Referent
Tertile 1 (1.3–42.0)	19,160	25	95,497	25	900	25	9.42	0.92	0.84, 1.01	0.97	0.88, 1.06
Tertile 2 (42.1–215.0)	18,916	24	94,984	24	757	21	7.97	0.74	0.68, 0.82	0.89	0.81, 0.98
Tertile 3 (215.1–1,000.0)	18,741	24	93,263	24	826	23	8.86	0.72	0.66, 0.79	0.89	0.81, 0.98
Missing data	597	1	2,927	1	44	1	15.03				
P-trend								<0.001		0.008	

Abbreviations: CI, confidence interval; HR, hazard ratio; N/A, not applicable.

[a] Because of rounding, numbers of person-years for each variable do not always sum to exactly 387,801.

[b] Number of deaths per 1,000 person-years.

[c] Adjusted for the following variables: sex; age; race/ethnicity; marital status; education; recency/dose of smoking; physical activity in the 10 years before baseline; estrogen therapy; estrogen plus progestin therapy; regular use of regular or extra-strength aspirin in the past 10 years; regular use of nonaspirin nonsteroidal antiinflammatory medication in the past 10 years; current use of cholesterol-lowering medication; prostate-specific antigen screening in the past 2 years; receipt of a mammogram in the past 2 years; sigmoidoscopy in the past 10 years; self-rated health; mother's and father's ages at death; body mass index at age 45 years; average alcohol intake at age 45 years; morbidity score; and the following variables, categorized in quartiles and a missing category: percentage of calories derived from *trans* fat; percentage of calories derived from saturated fat; number of servings per day of fruits; and number of servings per day of vegetables (excluding potatoes).

[d] *P*-trend is not applicable because the test for nonlinearity in the log hazard ratio was statistically significant at the 5% level.

[e] Of single supplements (and mixtures other than multivitamins).

[f] From single supplements (and mixtures other than multivitamins) plus multivitamins.

Although we adjusted for receipt of screening for several (but not all) cancers and for use of some medications that prevent CVD mortality, confounding by unmeasured healthy behaviors may have been present.

Table 3. Hazard Ratios for Total Mortality Associated With Use of Vitamin C and Vitamin E Supplements During the 10 Years Before Baseline, According to Participant Characteristics, Vitamins and Lifestyle Study, Western Washington State, 2000–2006

Characteristic	No. of Deaths in Reference Group (No Use)	Tertile 1			Tertile 2			Tertile 3		
		No. of Deaths	Multivariate-Adjusted HR[a]	95% CI	No. of Deaths	Multivariate-Adjusted HR[a]	95% CI	No. of Deaths	Multivariate-Adjusted HR[a]	95% CI
		Ten-Year Average Dose of Vitamin C[b], mg/day								
		2.6–60.0			60.1–322.0			322.1–1,750.0		
Sex										
Female	352	414	0.94	0.81, 1.10	378	1.02	0.87, 1.19	346	0.87	0.74, 1.03
Male	711	511	0.98	0.87, 1.11	406	0.91	0.80, 1.04	416	0.88	0.78, 1.01
P for interaction					0.685					
Smoking status										
Never smoker	301	275	0.91	0.77, 1.09	270	0.97	0.81, 1.16	211	0.76	0.63, 0.92
Former smoker; quit ≥10 years previously	404	381	1.04	0.90, 1.20	300	0.93	0.79, 1.09	322	0.91	0.77, 1.06
Current/recent smoker; quit <10 years previously	338	250	0.96	0.81, 1.15	188	1.00	0.83, 1.21	211	1.05	0.87, 1.27
P for interaction					0.143					
Body mass index[c] at age 45 years										
<25	479	445	0.96	0.84, 1.10	398	1.00	0.87, 1.15	409	0.94	0.81, 1.08
25–<30	326	271	1.04	0.88, 1.23	239	0.97	0.81, 1.17	219	0.91	0.76, 1.09
≥30	170	130	0.89	0.70, 1.14	93	0.87	0.67, 1.14	80	0.76	0.57, 1.01
P for interaction					0.771					
No. of servings of fruits and vegetables per day										
Less than median (0.0–3.1)	606	447	0.92	0.81, 1.05	348	0.94	0.82, 1.08	334	0.85	0.74, 0.98
Median or higher (3.2–26.4)	296	327	1.08	0.92, 1.28	322	1.02	0.87, 1.21	338	0.92	0.78, 1.09
P for interaction					0.740					
Morbidity score[d]										
No comorbid conditions	202	141	0.94	0.75, 1.18	129	0.91	0.71, 1.15	124	0.79	0.62, 1.01
≥1 comorbid condition	855	779	0.98	0.89, 1.09	649	0.98	0.80, 1.10	636	0.91	0.82, 1.02
P for interaction					0.238					
		Ten-Year Average Dose of Vitamin E[b], mg/day								
		2.3–42.0			42.1–215.0			215.1–1,000.0		
Sex										
Female	342	404	0.94	0.80, 1.09	345	0.91	0.78, 1.07	399	0.90	0.77, 1.05
Male	708	496	0.98	0.87, 1.11	412	0.86	0.76, 0.98	427	0.87	0.76, 0.99
P for interaction					0.474					
Smoking status										
Never smoker	287	271	0.92	0.77, 1.10	261	0.96	0.80, 1.15	235	0.80	0.66, 0.97
Former smoker; quit ≥10 years previously	398	344	0.96	0.83, 1.12	309	0.86	0.73, 1.00	360	0.89	0.77, 1.04
Current/recent smoker; quit <10 years previously	345	263	1.02	0.86, 1.21	166	0.83	0.68, 1.02	211	1.01	0.84, 1.22
P for interaction					0.570					

Body mass index at age 45 years											
<25	476	437	0.96	0.83, 1.10	364	0.87	0.75, 1.00	450	0.95	0.82, 1.09
25–<30	325	262	1.04	0.88, 1.23	236	0.96	0.80, 1.15	234	0.88	0.74, 1.06
≥30	162	126	0.85	0.67, 1.10	102	0.89	0.68, 1.16	83	0.78	0.58, 1.04
P for interaction					0.852					
No. of servings of fruits and vegetables per day										
Less than median (0.0–3.1)	609	446	0.94	0.82, 1.06	342	0.81	0.70, 0.93	337	0.82	0.71, 0.94
Median or higher (3.2–26.4)	286	298	1.00	0.84, 1.19	314	1.01	0.85, 1.20	314	1.00	0.85, 1.18
P for interaction					0.039					
Morbidity score[d]										
No comorbid conditions	205	143	0.89	0.71, 1.11	122	0.80	0.63, 1.02	128	0.86	0.68, 1.10
≥1 comorbid condition	839	749	0.99	0.89, 1.09	635	0.91	0.82, 1.02	694	0.90	0.81, 1.01
P for interaction					0.522					

Abbreviations: CI, confidence interval; HR, hazard ratio.
[a] Reference category, no use. See Table 2, footnote "c," for adjustment factors.
[b] From single supplements (and mixtures other than multivitamins) plus multivitamins.
[c] Weight (kg)/height (m)^2.
[d] See Table 1, footnote "c," for the list of comorbid conditions.

Further, although participants reported their use of supplements during the 10 years before baseline and were followed for mortality for 5 years, this etiologic time window may be too short for some diseases. Additionally, the sensitivity of this study to detect an association between use of multivitamins and mortality may have been low because of the fortification of enriched grain products with folic acid, which became mandatory in the United States in 1998 (31).

Another concern is exposure measurement error. Although we obtained detailed information on the duration, frequency, and daily dose of supplements used, these self-reported measures are subject to error. However, in a validity study (32) conducted in the Vitamins and Lifestyle Study cohort, the reliability and validity of the measures of supplement use were found to be quite good. For the variable 10-year average dose, the intraclass correlation coefficient for test-retest reliability at baseline and after 3 months was 0.81 for multivitamins, 0.85 for vitamin C, and 0.87 for vitamin E. As compared with an interviewer's transcription of nutrient information on bottle labels, Pearson's correlation coefficient was 0.77 for current use of vitamin C and 0.81 for current use of vitamin E. As compared with vitamin nutrient levels in the blood, Pearson's correlation coefficient was 0.29 for intake of vitamin C from supplements and 0.69 for intake of vitamin E from supplements.

Below we compare our findings with findings from prior cohort studies and randomized trials of these associations.

Total mortality

Multivitamins. Our finding of no association between use of multivitamins and total mortality is consistent with the 2 prior cohort studies of this relation (10, 33). Although there are no published results from randomized trials of the common formulations of multivitamins, in 2 randomized trials of combinations of vitamins and minerals (the Linxian Trials (34) and the SU.VI.MAX Study (35)), small inverse associations were observed (relative risk (RR) = 0.87 (34) and RR = 0.77 (35)).

Vitamin C. Our finding of a decreased risk of total mortality associated with use of vitamin C supplements is consistent with some (11, 16) but not all (10) cohort studies; reported relative risks range from 0.85 to 1.09 (10, 11, 16). In a meta-analysis of 3 randomized trials of vitamin C supplement use and total mortality, the summary relative risk was 0.88 (95% CI: 0.32, 2.42) (20), and recent findings from 2 large randomized trials that were not included in the meta-analysis do not support an association. In one, the Women's Antioxidant Cardiovascular Study (WACS), which was conducted among 8,171 female health professional at elevated risk for cardiovascular events, the relative risk of total mortality associated with vitamin C supplement use (500 mg/day) was 1.03 during a mean follow-up period of 9 years (36). In the other, Physicians' Health Study (PHS) II, which was conducted among 16,641 male health professionals, the corresponding hazard ratio (500 mg of vitamin C per day) was 1.07 during a mean follow-up period of 8 years (37).

Vitamin E. Our finding of a decreased risk of total mortality is consistent with most (9, 10, 15, 17) but not all (11)

Table 4. Hazard Ratios for Cardiovascular Disease Mortality, Cancer Mortality, and Mortality From All Other Causes Combined Associated With Use of Vitamin Supplements During the 10 Years Before Baseline, Vitamins and Lifestyle Study, Western Washington State, 2000–2006

Cause of Death	No. of Deaths in Reference Group (No Use)	Tertile 1			Tertile 2			Tertile 3			*P* for Trend
		No. of Deaths	Multivariate-Adjusted HR[a]	95% CI	No. of Deaths	Multivariate-Adjusted HR[a]	95% CI	No. of Deaths	Multivariate-Adjusted HR[a]	95% CI	
		Ten-Year Average Frequency of Multivitamin Use, days/week									
		>0–2			3–5			6–7			
Cardiovascular disease											
Total[b]	350	140	1.00	0.81, 1.24	94	0.78	0.61, 1.00	285	0.84	0.70, 0.99	0.019
No history[c]	200	85	1.08	0.82, 1.41	51	0.72	0.52, 1.01	152	0.78	0.62, 0.98	0.012
History[c]	150	55	0.88	0.62, 1.24	43	0.84	0.58, 1.22	133	0.92	0.71, 1.19	0.498
Cancer											
Total[d]	578	271	1.07	0.92, 1.25	206	1.00	0.84, 1.18	609	1.06	0.94, 1.20	0.415
No history[e]	292	129	1.13	0.91, 1.41	104	1.05	0.82, 1.33	303	1.07	0.90, 1.28	0.517
History[e]	286	142	1.01	0.81, 1.23	102	0.94	0.74, 1.20	306	1.04	0.87, 1.25	0.705
All other causes (total)[f]	320	132	1.00	0.80, 1.23	105	0.93	0.73, 1.18	339	1.12	0.95, 1.32	0.233
		Ten-Year Average Dose of Vitamin C[g], mg/day									
		2.6–60.0			60.1–322.0			322.1–1,750.0			
Cardiovascular disease											
Total[b]	293	215	0.89	0.74, 1.08	179	0.81	0.66, 0.99	201	0.89	0.73, 1.08	0.147
No history[c]	165	114	0.82	0.64, 1.06	99	0.84	0.64, 1.10	120	1.00	0.77, 1.29	0.998
History[c]	128	101	0.95	0.72, 1.27	80	0.79	0.57, 1.08	81	0.75	0.55, 1.02	0.036
Cancer											
Total[d]	487	451	1.00	0.87, 1.15	399	1.07	0.93, 1.23	349	0.84	0.73, 0.98	N/A[h]
No history[e]	254	228	1.06	0.87, 1.28	175	0.90	0.74, 1.11	179	0.86	0.69, 1.05	0.076
History[e]	233	223	0.95	0.78, 1.16	224	1.20	0.98, 1.46	170	0.82	0.66, 1.02	N/A[h]
All other causes (total)[f]	267	244	1.03	0.86, 1.24	194	1.00	0.82, 1.22	200	1.01	0.82, 1.23	0.984
		Ten-Year Average Dose of Vitamin E[g], mg/day									
		1.3–42.0			42.1–215.0			215.1–1,000.0			
Cardiovascular disease											
Total[b]	300	201	0.79	0.64, 0.96	185	0.75	0.61, 0.92	203	0.72	0.59, 0.88	0.001
No history[c]	174	117	0.77	0.60, 0.99	107	0.79	0.61, 1.03	103	0.70	0.53, 0.91	0.012
History[c]	126	84	0.80	0.59, 1.09	78	0.70	0.51, 0.96	100	0.73	0.54, 0.99	0.027
Cancer											
Total[d]	476	445	1.02	0.89, 1.17	357	0.92	0.79, 1.06	410	0.93	0.81, 1.08	0.206
No history[e]	248	210	1.00	0.82, 1.22	172	0.86	0.70, 1.06	211	0.96	0.78, 1.17	0.422
History[e]	228	235	1.03	0.84, 1.25	185	0.95	0.77, 1.17	199	0.89	0.73, 1.11	0.235
All other causes (total)[f]	258	240	1.08	0.90, 1.31	204	0.99	0.81, 1.21	200	0.98	0.80, 1.20	0.679

Abbreviations: CI, confidence interval; HR, hazard ratio.

[a] Reference category, no use.

[b] HR was adjusted for a history of cardiovascular disease at baseline (defined as a previous heart attack, coronary bypass surgery, angioplasty, a diagnosis of angina, or a previous stroke); use of blood pressure medication in the past 2 weeks; history of heart attack in a first-degree relative; and diabetes at baseline, in addition to all of the variables listed in footnote "c" of Table 2, except morbidity score, mother's age at death, and father's age at death.

[c] HR was adjusted for all of the variables specified in the above footnote within strata of history of cardiovascular disease.

[d] HR was adjusted for history of cancer and family history of cancer, in addition to the variables listed in footnote "c" of Table 2, except morbidity score, mother's age at death, and father's age at death.

[e] HR was adjusted for all of the variables specified in the above footnote within strata of history of cancer.

[f] HR was adjusted for all of the variables listed in footnote "c" of Table 2.

[g] From single supplements (and mixtures other than multivitamins) and multivitamins.

[h] *P*-trend is not applicable because the test for nonlinearity in the log hazard ratio was statistically significant at the 5% level.

cohort studies. Reported relative risks range from 0.73 to 1.44 (9–11, 15, 17); however, randomized trials do not support an association. In a meta-analysis of 24 randomized trials of vitamin E supplement use and total mortality, the summary relative risk was 1.02 (95% CI: 0.98, 1.05) (20). The relative risk of total mortality associated with use of vitamin E supplements was 1.00 in WACS (600 IU of vitamin E every other day) (36), and the hazard ratio was 1.07 in PHS II (400 IU of vitamin E every other day) (37).

Mortality from CVD and cancer

Multivitamins. Although in the present study there was a slightly decreased risk of CVD mortality associated with use of multivitamins, results from a prior cohort study suggested no association with coronary heart disease mortality (10). Risk of cerebrovascular disease mortality was evaluated in the Linxian Trials; the relative risk was 0.90 (34).

Our finding of no association between use of multivitamins and risk of cancer mortality is consistent with findings from a prior cohort study (10). The relative risk of cancer mortality was 0.87 in the Linxian Trials (34).

Vitamin C. Use of vitamin C supplements was associated with a decreased risk of coronary heart disease mortality in a 2004 pooled analysis of data from 4 cohort studies (RR = 0.76, 95% CI: 0.58, 0.99) (12). In the present study, overall, use of vitamin C was not associated with CVD mortality, although there was a slightly decreased risk among persons with a history of CVD at baseline. The association between vitamin C supplement use and CVD mortality was evaluated in 2 randomized trials (36, 37); the relative risk was 1.10 in WACS (36), and the hazard ratio was 1.02 in PHS II (37).

In the present study, use of vitamin C was associated with a slightly decreased risk of cancer mortality, although there was no dose-response trend. In a prior cohort study carried out among the elderly, the relative risk was 0.88 (10). However, no inverse association was observed in 2 randomized trials of the association between vitamin C use and cancer mortality; the relative risk was 1.28 in WACS (38), and the hazard ratio was 1.06 in PHS II (39).

Vitamin E. In the present study, use of vitamin E was associated with a decreased risk of CVD mortality. In a 2004 meta-analysis of results from 7 randomized trials on the association between vitamin E use and CVD mortality, Eidelman et al. (21) found a summary relative risk of 1.00 (95% CI: 0.94, 1.05). This finding is consistent with findings from another 2004 meta-analysis of 5 randomized trials (4 were included in the Eidelman et al. study) (18) and a pooled analysis of 4 cohort studies of coronary heart disease mortality (12).

Since 2004, there have been additional randomized trials. In the 2005 Women's Health Study, which included 39,876 women, risk of CVD mortality was lower among women in the vitamin E arm (600 IU every other day) relative to the placebo arm during a mean follow-up period of 10 years (RR = 0.76) (40). The duration of the vitamin E intervention was longer in that study than in previous trials (40). However, since the publication of the Women's Health Study results, vitamin E has been found to not be associated with CVD mortality in PHS II (RR = 1.07) (37) or WACS (hazard

ratio = 0.94) (36), and both of those studies had treatment durations almost as long as that of the Women's Health Study.

Use of vitamin E supplements was not associated with cancer mortality in the present study. This finding is consistent with that observed in a meta-analysis of 4 randomized trials (22). The relative risk was 0.87 in WACS (38), and the hazard ratio was 1.13 in PHS II (39). In a prior cohort study of elderly persons, the relative risk was 0.81 (10).

Associations stratified by potential modifiers

Our findings of stronger associations between total mortality risk and use of vitamins C and E among persons with greater body mass index and lesser fruit and vegetable consumption are consistent with the hypothesis that any impact of vitamins C and E on total mortality risk may be stronger among persons with greater levels of oxidative stress (1). On the other hand, the associations between use of vitamins C and E and total mortality risk were stronger among never smokers than among current/recent smokers, yet smoking is thought to increase oxidative stress (1). Notably, in a separate study, vitamin C and E supplements were associated with increased risks of total mortality among smokers but not among nonsmokers (13).

Summary

In the present study, we observed small decreased risks of total mortality associated with use of vitamin C and E supplements, but we found no association with multivitamins. In cause-specific analyses, multivitamin use and vitamin E use were associated with decreased risks of CVD mortality. Although the association between vitamin E use and CVD mortality was consistent with that observed in the Women's Health Study randomized trial, other findings were small in magnitude and should be interpreted cautiously because healthy behaviors tend to be more common in supplement users than in nonusers.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington (Gaia Pocobelli, Ulrike Peters, Alan R. Kristal, Emily White); and Cancer Prevention Program, Fred Hutchinson Cancer Research Center, Seattle, Washington (Gaia Pocobelli, Ulrike Peters, Alan R. Kristal, Emily White).

This research was funded through National Institutes of Health grant R01 CA74846 and a grant from the Washington State Vitamins Distribution Agreement.

Conflict of interest: none declared.

REFERENCES

1. Halliwell B, Gutteridge JM. *Free Radicals in Biology and Medicine*. 4th ed. New York, NY: Oxford University Press; 2007.
2. Young IS, Woodside JV. Antioxidants in health and disease. *J Clin Pathol*. 2001;54(3):176–186.
3. Loft S, Møller P, Cooke MS, et al. Antioxidant vitamins and cancer risk: is oxidative damage to DNA a relevant biomarker? *Eur J Nutr.* 2008;47(suppl 2):19–28.
4. Dizdaroglu M. Oxidative damage to DNA in mammalian chromatin. *Mutat Res.* 1992;275(3–6):331–342.
5. Dragsted LO. Biomarkers of exposure to vitamins A, C, and E and their relation to lipid and protein oxidation markers. *Eur J Nutr.* 2008;47(suppl 2):3–18.
6. Weinberg RA. How cancer arises. *Sci Am.* 1996;275(3):62–70.
7. Steinberg D, Lewis A. Conner Memorial Lecture. Oxidative modification of LDL and atherogenesis. *Circulation*. 1997; 95(4):1062–1071.
8. Qato DM, Alexander GC, Conti RM, et al. Use of prescription and over-the-counter medications and dietary supplements among older adults in the United States. *JAMA*. 2008;300(24): 2867–2878.
9. Stampfer MJ, Hennekens CH, Manson JE, et al. Vitamin E consumption and the risk of coronary disease in women. *N Engl J Med*. 1993;328(20):1444–1449.
10. Losonczy KG, Harris TB, Havlik RJ. Vitamin E and vitamin C supplement use and risk of all-cause and coronary heart disease mortality in older persons: the Established Populations for Epidemiologic Studies of the Elderly. *Am J Clin Nutr.* 1996;64(2):190–196.
11. Sahyoun NR, Jacques PF, Russell RM. Carotenoids, vitamins C and E, and mortality in an elderly population. *Am J Epidemiol*. 1996;144(5):501–511.
12. Knekt P, Ritz J, Pereira MA, et al. Antioxidant vitamins and coronary heart disease risk: a pooled analysis of 9 cohorts. *Am J Clin Nutr.* 2004;80(6):1508–1520.
13. Brzozowska A, Kaluza J, Knoops KT, et al. Supplement use and mortality: the SENECA Study. *Eur J Nutr.* 2008;47(3): 131–137.
14. Messerer M, Håkansson N, Wolk A, et al. Dietary supplement use and mortality in a cohort of Swedish men. *Br J Nutr.* 2008; 99(3):626–631.
15. Hayden KM, Welsh-Bohmer KA, Wengreen HJ, et al. Risk of mortality with vitamin E supplements: the Cache County Study. *Am J Med*. 2007;120(2):180–184.
16. Enstrom JE, Kanim LE, Klein MA. Vitamin C intake and mortality among a sample of the United States population. *Epidemiology.* 1992;3(3):194–202.
17. Rimm EB, Stampfer MJ, Ascherio A, et al. Vitamin E consumption and the risk of coronary heart disease in men. *N Engl J Med.* 1993;328(20):1450–1456.
18. Shekelle PG, Morton SC, Jungvig LK, et al. Effect of supplemental vitamin E for the prevention and treatment of cardiovascular disease. *J Gen Intern Med*. 2004;19(4): 380–389.
19. Miller ER III, Pastor-Barriuso R, Dalal D, et al. Meta-analysis: high-dosage vitamin E supplementation may increase all-cause mortality. *Ann Intern Med*. 2005;142(1):37–46.
20. Bjelakovic G, Nikolova D, Gluud LL, et al. Antioxidant supplements for prevention of mortality in healthy participants and patients with various diseases [electronic article]. *Cochrane Database Syst Rev.* 2008;(2):CD007176.
21. Eidelman RS, Hollar D, Hebert PR, et al. Randomized trials of vitamin E in the treatment and prevention of cardiovascular disease. *Arch Intern Med*. 2004;164(14):1552–1556.
22. Bardia A, Tleyjeh IM, Cerhan JR, et al. Efficacy of antioxidant supplementation in reducing primary cancer incidence and mortality: systematic review and meta-analysis. *Mayo Clin Proc.* 2008;83(1):23–34.

23. Kristal AR. Are clinical trials the ''gold standard'' for cancer prevention research? *Cancer Epidemiol Biomarkers Prev.* 2008;17(12):3289–3291.
24. White E, Patterson RE, Kristal AR, et al. VITamins And Lifestyle cohort study: study design and characteristics of supplement users. *Am J Epidemiol*. 2004;159(1):83–93.
25. Medical Economics Company. *Physicians' Desk Reference for Nonprescription Drugs and Dietary Supplements 2002*. Montvale, NJ: Medical Economics Company; 2002.
26. Littman AJ, White E, Kristal AR, et al. Assessment of a one-page questionnaire on long-term recreational physical activity. *Epidemiology.* 2004;15(1):105–113.
27. Patterson RE, Kristal AR, Tinker LF, et al. Measurement characteristics of the Women's Health Initiative food frequency questionnaire. *Ann Epidemiol*. 1999;9(3):178–187.
28. US Department of Health and Human Services and US Department of Agriculture. *Dietary Guidelines for Americans*. Washington, DC: US GPO; 2005.
29. Cox DR, Oakes D. *Analysis of Survival Data*. London, United Kingdom: Chapman & Hall Ltd; 1984.
30. World Health Organization. *International Statistical Classification of Diseases and Related Health Problems, 10th Revision. Version for 2007* [electronic book]. Geneva, Switzerland: World Health Organization; 2007. (http://www.who.int/classifications/apps/icd/icd10online/). (Accessed October 13, 2008).
31. Quinlivan EP, Gregory JH III. Effect of food fortification on folic acid intake in the United States. *Am J Clin Nutr.* 2003; 77(1):221–225.
32. Satia-Abouta J, Patterson RE, King IB, et al. Reliability and validity of self-report of vitamin and mineral supplement use in the Vitamins and Lifestyle Study. *Am J Epidemiol*. 2003; 157(10):944–954.
33. Watkins ML, Erickson JD, Thun MJ, et al. Multivitamin use and mortality in a large prospective study. *Am J Epidemiol*. 2000;152(2):149–162.
34. Blot WJ, Li JY, Taylor PR, et al. Nutrition intervention trials in Linxian, China: supplementation with specific vitamin/mineral combinations, cancer incidence, and disease-specific mortality in the general population. *J Natl Cancer Inst*. 1993; 85(18):1483–1492.
35. Hercberg S, Galan P, Preziosi P, et al. The SU.VI.MAX Study: a randomized, placebo-controlled trial of the health effects of antioxidant vitamins and minerals. *Arch Intern Med.* 2004; 164(21):2335–2342.
36. Cook NR, Albert CM, Gaziano JM, et al. A randomized factorial trial of vitamins C and E and beta carotene in the secondary prevention of cardiovascular events in women: results from the Women's Antioxidant Cardiovascular Study. *Arch Intern Med.* 2007;167(15):1610–1618.
37. Sesso HD, Buring JE, Christen WG, et al. Vitamins E and C in the prevention of cardiovascular disease in men: the Physicians' Health Study II randomized controlled trial. *JAMA*. 2008;300(18):2123–2133.
38. Lin J, Cook NR, Albert C, et al. Vitamins C and E and beta carotene supplementation and cancer risk: a randomized controlled trial. *J Natl Cancer Inst*. 2009;101(1):14–23.
39. Gaziano JM, Glynn RJ, Christen WG, et al. Vitamins E and C in the prevention of prostate and total cancer in men: the Physicians' Health Study II randomized controlled trial. *JAMA*. 2009;301(1):52–62.
40. Lee IM, Cook NR, Gaziano JM, et al. Vitamin E in the primary prevention of cardiovascular disease and cancer: the Women's Health Study: a randomized controlled trial. *JAMA*. 2005; 294(1):56–65.

Vitamin E in the Primary Prevention of Cardiovascular Disease and Cancer: The Women's Health Study: A Randomized Controlled Trial

I-Min Lee; Nancy R. Cook; J. Michael Gaziano; et al.

JAMA. 2005;294(1):56-65 (doi:10.1001/jama.294.1.56)

http://jama.ama-assn.org/cgi/content/full/294/1/56

Vitamin E in the Primary Prevention of Cardiovascular Disease and Cancer

The Women's Health Study: A Randomized Controlled Trial

I-Min Lee, MBBS, ScD
Nancy R. Cook, ScD
J. Michael Gaziano, MD
David Gordon, MA
Paul M Ridker, MD
JoAnn E. Manson, MD, DrPH
Charles H. Hennekens, MD, DrPH
Julie E. Buring, ScD

Context Basic research provides plausible mechanisms and observational studies suggest that apparently healthy persons, who self-select for high intakes of vitamin E through diet or supplements, have decreased risks of cardiovascular disease and cancer. Randomized trials do not generally support benefits of vitamin E, but there are few trials of long duration among initially healthy persons.

Objective To test whether vitamin E supplementation decreases risks of cardiovascular disease and cancer among healthy women.

Design, Setting, and Participants In the Women's Health Study conducted between 1992 and 2004, 39 876 apparently healthy US women aged at least 45 years were randomly assigned to receive vitamin E or placebo and aspirin or placebo, using a 2×2 factorial design, and were followed up for an average of 10.1 years.

Intervention Administration of 600 IU of natural-source vitamin E on alternate days.

Main Outcome Measures Primary outcomes were a composite end point of first major cardiovascular event (nonfatal myocardial infarction, nonfatal stroke, or cardiovascular death) and total invasive cancer.

Results During follow-up, there were 482 major cardiovascular events in the vitamin E group and 517 in the placebo group, a nonsignificant 7% risk reduction (relative risk [RR], 0.93; 95% confidence interval [CI], 0.82-1.05; *P*=.26). There were no significant effects on the incidences of myocardial infarction (RR, 1.01; 95% CI, 0.82-1.23; *P*=.96) or stroke (RR, 0.98; 95% CI, 0.82-1.17; *P*=.82), as well as ischemic or hemorrhagic stroke. For cardiovascular death, there was a significant 24% reduction (RR, 0.76; 95% CI, 0.59-0.98; *P*=.03). There was no significant effect on the incidences of total cancer (1437 cases in the vitamin E group and 1428 in the placebo group; RR, 1.01; 95% CI, 0.94-1.08; *P*=.87) or breast (RR, 1.00; 95% CI, 0.90-1.12; *P*=.95), lung (RR, 1.09; 95% CI, 0.83-1.44; *P*=.52), or colon cancers (RR, 1.00; 95% CI, 0.77-1.31; *P*=.99). Cancer deaths also did not differ significantly between groups. There was no significant effect of vitamin E on total mortality (636 in the vitamin E group and 615 in the placebo group; RR, 1.04; 95% CI, 0.93-1.16; *P*=.53).

Conclusions The data from this large trial indicated that 600 IU of natural-source vitamin E taken every other day provided no overall benefit for major cardiovascular events or cancer, did not affect total mortality, and decreased cardiovascular mortality in healthy women. These data do not support recommending vitamin E supplementation for cardiovascular disease or cancer prevention among healthy women.

JAMA. 2005;294:56-65 www.jama.com

FREE RADICALS CAN CAUSE LIPID peroxidation and DNA damage, contributing to the development of cardiovascular disease (CVD) and cancer.[1-5] Vitamin E has antioxidant properties, including inhibition of oxidation of low-density lipoprotein cholesterol in plasma, leading to the hypothesis that it can prevent these chronic diseases.[5] In some, but not all, basic research reports, vitamin E supplementation retarded atherogenesis.[6] In descriptive data, investigators noted a strong inverse relation between plasma vitamin E concentrations and death rates from ischemic heart disease in men in several European countries.[7] Additionally, several large cohort studies observed decreased CVD rates among individuals who self-selected for higher intakes of vitamin E through diet and/or supplements.[8-10] By 1997, despite a lack of randomized trials, 44% of US cardiologists reported routine use of antioxidant supplements, primarily vitamin E, compared with 42% who routinely used aspirin for the primary prevention of CVD.[11]

See also pp 47 and 105.

Author Affiliations are listed at the end of this article.
Corresponding Author: I-Min Lee, MBBS, ScD, Brigham and Women's Hospital, 900 Commonwealth Ave E, Boston, MA 02215 (ilee@rics.bwh.harvard.edu).

With regard to cancer, several observational studies, particularly case-control studies, also reported reduced rates of cancer among persons who self-selected for high antioxidant intakes.[12]

For small to moderate effects, however, the amount of uncontrolled and uncontrollable confounding inherent in observational studies can be as large as the postulated benefit, so randomized clinical trials represent the most reliable study design strategy.[13] Several trials were therefore initiated beginning in the late 1980s to directly test the vitamin E hypothesis.[14-38] To date, data from randomized trials have largely demonstrated no significant benefit of vitamin E supplementation on the incidence of CVD or cancer and, indeed, raised the question of possible adverse effects on total mortality with high doses.[39-44] However, these trials have been conducted primarily among participants with cardiovascular risk factors and/or CVD or at high risk for cancer. Few trials have recruited apparently healthy persons, with most designed to examine ophthalmologic outcomes.[30,33,37] Only one trial, testing a combination of antioxidant vitamins and minerals, has investigated CVD and cancer prevention among healthy persons not selected based on risk factors.[35] Additionally, the treatment duration in previous trials has generally been limited to 5 years or shorter, with 6 trials having a longer duration.[14-16,30,35,38] One possible explanation for the largely null results of randomized trials is that the duration of supplementation has been insufficient for an effect.[45]

To provide further information, the Women's Health Study (WHS) tested whether vitamin E supplementation for 10 years decreased risks of CVD and cancer in a large group of healthy women.

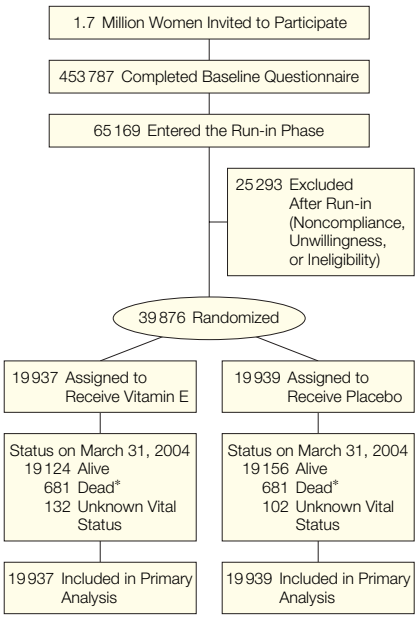
METHODS

Study Design

The WHS was a randomized, double-blind, placebo-controlled, 2×2 factorial trial evaluating the balance of risks and benefits of low-dose aspirin (100 mg every other day; Bayer Healthcare) and vitamin E (600 IU of α-tocopherol every other day; Natural Source Vitamin E Association) in the primary prevention of CVD and cancer.[46,47] Originally, a third component, beta carotene, was also included. However, this component was terminated early in January 1996 after a median treatment duration of 2.1 years.[48] Written informed consent was obtained from all participating women. The trial was approved by the institutional review board of Brigham and Women's Hospital and monitored by an external data and safety monitoring board.

Detailed methods of the design have been described previously.[46,47] Briefly, between September 1992 and May 1995, letters of invitation to participate in the trial and baseline health questionnaires were mailed to more than 1.7 million female health care professionals throughout the United States (FIGURE 1). A total of 453 787 women completed the questionnaires and 65 169 were willing and eligible to participate. Eligibility criteria included the following: age 45 years or older; no previous history of coronary heart disease, cerebrovascular disease, cancer (except nonmelanoma skin cancer), or other major chronic illnesses; no history of adverse effects from aspirin; no use of aspirin or nonsteroidal anti-inflammatory drugs (NSAIDs) more than once a week, or willingness to forgo their use; no use of anticoagulants or corticosteroids; and no use of individual supplements of vitamin A, E, or beta carotene more than once a week. Eligible women were enrolled into a 3-month run-in period with placebo medications to identify likely long-term compliers to pill taking. Following the run-in period, 39 876 women remained willing, eligible, and compliant, and they were randomized in blocks of 16 within 5-year age strata to vitamin E (n=19 937) or placebo (n=19 939).

Figure 1. Flow Diagram of the Vitamin E Component of the Women's Health Study



*The numbers of deaths are higher than those in Table 2 because this figure includes all reported deaths in the WHS, whereas the deaths in Table 2 include only reported deaths confirmed by the Endpoints Committee or by a death certificate.

Study Treatment and Follow-up

Each year, women received calendar packs that contained amber capsules (vitamin E or placebo) and white pills (aspirin or placebo) on alternate days. Every 6 months for the first year and annually thereafter, they also received follow-up questionnaires inquiring about compliance with pill-taking, potential adverse effects, occurrence of end points, and risk factors. Study medications and end point ascertainment were continued in blinded fashion through the scheduled end of the trial (March 31, 2004). Follow-up and validation of reported end points were completed in February 2005. Morbidity and mortality follow-up were 97.2% and 99.4% complete, respectively.

Using the information provided on questionnaires, compliance, defined as taking at least two thirds of the study capsules, was 78.9% at 5 years and 71.6% at 10 years. Averaged throughout the trial, it was 75.8% with no difference between active and placebo groups ($P=.64$). Nontrial use of individual supplements of vitamin E for at least 4 days per month ("drop-ins") was

10.0% at 5 years and 10.9% at 10 years. Averaged throughout the trial, outside use was somewhat lower in the active (8.6%) than in the placebo group (8.9%) ($P=.07$).

Confirmation of End Points

The primary end points were a composite of first major cardiovascular event (nonfatal myocardial infarction [MI], nonfatal stroke, or cardiovascular death) and total invasive cancer (apart from nonmelanoma skin cancer). Secondary end points were the individual cardiovascular events—total MI, total stroke, and cardiovascular death—and the main site-specific cancers in women: breast, lung, and colon cancers. We also collected information on coronary revascularization procedures (bypass surgery or percutaneous coronary angioplasty), transient ischemic attacks (TIAs), and total mortality.

Women reported the occurrence of relevant end points via follow-up questionnaires, letters, or telephone calls. Deaths were usually reported by family members or postal authorities or ascertained through the National Death Index. After obtaining written consent, we acquired medical records from hospitals and physicians, which were reviewed by the WHS Endpoints Committee of physicians blinded to randomized treatment assignment. The committee confirmed a diagnosis of MI if symptoms met World Health Organization criteria and the event was associated with abnormal levels of cardiac enzymes or diagnostic electrocardiograms. The use of coronary revascularization procedures was confirmed by medical record review. Stroke was confirmed if the participant had a new neurologic deficit of sudden onset that persisted for more than 24 hours or until death within 24 hours. Clinical information and computed tomographic scans or magnetic resonance images were used to distinguish hemorrhagic from ischemic strokes.[49] A confirmed TIA was defined as a neurologic deficit of sudden onset lasting less than 24 hours. Cardiovascular deaths were confirmed by autopsy reports, death certificates, medical records, and information from next of kin or family members. The vast majority (96.8%) of cancers were confirmed with pathology or cytology reports. Rarely, the committee confirmed a reported case of cancer based on strong clinical and radiological or laboratory marker evidence (eg, elevated CA-125) when pathology or cytology review was not conducted. Total mortality was confirmed by the committee or by obtaining a death certificate. Only confirmed end points are included in this report.

Statistical Analysis

All primary analyses were performed on an intention-to-treat basis (ie, based on all randomized persons, as randomized), using the SAS statistical software package (release 8.2; SAS Institute Inc, Cary, NC). We used Cox proportional hazards regression models to calculate the relative risks (RRs) and 95% confidence intervals (95% CIs), comparing event rates in the vitamin E and placebo groups, after adjustment for age and randomized aspirin and beta carotene assignments. Statistical significance was set at $P<.05$, using 2-sided tests. To test the proportionality assumption (ie, that of nonchanging RRs over time), we included an interaction term of vitamin E with the logarithm of time in the Cox models. The proportionality assumption was not violated for major cardiovascular events ($P=.16$), total invasive cancer ($P=.72$), or total mortality ($P=.81$). We conducted subgroup analyses stratified by major risk factors for CVD and cancer, and assessed effect modification using interaction terms between subgroup indicators and vitamin E assignment, testing for trend when subgroup categories were ordinal. To investigate the effect of compliance, we carried out a sensitivity analysis that censored follow-up for any participant at the time when she reported taking less than two thirds of study medications over the previous year.

RESULTS

The mean (SD) age of participants at baseline was 54.6 (7.0) years; other clinical characteristics are shown in TABLE 1. As expected in this very large sample, randomization was effective in balancing the characteristics of women in the vitamin E and placebo groups. The average duration of follow-up from randomization to the end of the trial was 10.1 years (range, 8.2-10.9 years).

Cardiovascular Disease

By the end of the trial, 999 major cardiovascular events (253 per 100 000 person-years) had occurred: 482 in the vitamin E group and 517 in the placebo group (TABLE 2). This corresponded to a nonsignificant 7% risk reduction with vitamin E (RR, 0.93; 95% CI, 0.82-1.05; $P=.26$). For the individual cardiovascular events, vitamin E had no effect on total MI (RR, 1.01; 95% CI, 0.82-1.23) or total stroke (RR, 0.98; 95% CI, 0.82-1.17). For stroke subtypes, there was no reduction in ischemic or increase in hemorrhagic stroke rates. There was a significant 24% reduction in cardiovascular deaths among women in the vitamin E group (RR, 0.76; 95% CI, 0.59-0.98). This was largely attributable to fewer sudden deaths in the vitamin E group (38 vs 51 among women assigned to placebo) and fewer deaths from other cardiovascular disease (ie, deaths due to cardiovascular diseases other than ischemic heart disease and cerebrovascular disease, 20 vs 34, respectively). There was no significant effect of vitamin E on coronary revascularization procedures (394 vs 369, respectively) or TIA (212 in each group).

FIGURE 2 shows the cumulative incidence rates of major cardiovascular events among women in the 2 groups by year of follow-up. An apparent benefit of vitamin E on major cardiovascular events, as well as on the individual end points of MI, stroke, and cardiovascular death, was observed early in the trial. The effect on major cardiovascular events diminished over time and disappeared for MI and stroke by the end of the trial. In contrast, the difference in cardiovascular death rates between active and placebo groups appeared to increase over time; however, the change in RRs over time was not significant ($P=.59$). Because com-

pliance diminishes over time, we examined whether the observed trends might have been due to this tendency. In a sensitivity analysis that censored noncompliant (taking less than two thirds of study medications) follow-up time, there was no evidence that noncompliance influenced the findings (RR for major cardiovascular events, 0.96; 95% CI, 0.82-1.11; $P=.56$).

As reported previously, aspirin was associated with a nonsignificant 9% reduction in major cardiovascular events.[47] We therefore examined whether random assignment to aspirin modified the effect of vitamin E. There was no modification of the effect of vitamin E by random assignment to aspirin (TABLE 3). Beta carotene also did not modify the effect of vitamin E on the primary or secondary end points (data not shown).

We examined whether cardiovascular risk factors modified the relation between vitamin E and major cardiovascular events (Table 3). In particular, we were interested in whether levels of oxidative stress may modify the effect of vitamin E.[50] We did not have a direct measure of oxidative stress; however, smoking and diseases such as hypertension, hyperlipidemia, and diabetes are associated with increased production of reactive oxygen species in the vascular wall.[51] Using these indirect markers (all self-reported), we found no evidence of benefit of vitamin E among persons with increased oxidative stress. Additionally, no benefit was observed among both users and nonusers of multivitamins, who would presumably have lower and higher levels of oxidative stress, respectively.

There also was no statistically significant effect modification by any of the other factors considered, except age ($P=.008$). In subgroup analyses, women aged at least 65 years comprised 10% of study participants but contributed 31% of end points. A significant 26% reduction in major cardiovascular events was observed among women aged at least 65 years assigned to vitamin E (RR, 0.74; 95% CI, 0.59-0.93; $P=.009$) due to a 34% reduction in MI (RR, 0.66; 95% CI, 0.45-0.98; $P=.04$) and 49% reduction in cardiovascular death (RR, 0.51; 95% CI, 0.33-0.77; $P<.001$) rates. However, no reduction in stroke rate was observed (RR, 0.88; 95% CI, 0.64-1.21; $P=.44$). Among women aged 45 through 54 and 55 through 64 years, the RRs for major cardiovascular events were 1.13 (95% CI, 0.91-1.41; $P=.26$) and 0.95 (95% CI, 0.77-1.16; $P=.61$), respectively.

Cancer

During the trial, 2865 women developed invasive cancer (741 events per

Table 1. Baseline Characteristics of Women by Group, Women's Health Study

Characteristic	No. (%)* Vitamin E (n = 19 937)	Placebo (n = 19 939)	Total (N = 39 876)	*P* Value
Age, y				
Mean (SD)	54.6 (7.0)	54.6 (7.0)	54.6 (7.0)	.94
45-54	12 016 (60.3)	12 009 (60.2)	24 025 (60.2)	.98
55-64	5878 (29.5)	5876 (29.5)	11 754 (29.5)	
≥65	2043 (10.3)	2054 (10.3)	4097 (10.3)	
Smoking status				
Current	2590 (13.0)	2645 (13.3)	5235 (13.1)	.42
Past or never	17 328 (87.0)	17 277 (86.7)	34 605 (86.9)	
Alcohol intake				
Never/rarely	9057 (45.4)	8925 (44.8)	17 982 (45.1)	.18
At least 1/mo	10 873 (54.6)	11 011 (55.2)	21 884 (54.9)	
Multivitamin use	7807 (39.2)	7661 (38.4)	15 468 (38.8)	.13
Body mass index†				
Mean (SD)	26.04 (5.07)	26.03 (5.06)	26.04 (5.06)	.94
<25	9885 (50.7)	9964 (51.0)	19 849 (50.8)	.75
25-29	6069 (31.1)	6012 (30.8)	12 081 (30.9)	
≥30	3557 (18.2)	3569 (18.3)	7126 (18.2)	
Physical activity, kcal/wk				
<1000	13 030 (66.2)	12 964 (65.8)	25 994 (66.0)	.37
≥1000	6645 (33.8)	6738 (34.2)	13 383 (34.0)	
Menopausal status and hormone therapy use				.88
Premenopausal	5458 (27.5)	5515 (27.7)	10 973 (27.6)	
Uncertain	3568 (17.9)	3581 (18.0)	7149 (18.0)	
Postmenopausal, current hormone therapy use	5981 (30.1)	5967 (30.0)	11 948 (30.0)	
Postmenopausal, never or past hormone therapy use	4880 (24.5)	4824 (24.3)	9704 (24.4)	
Hypertension‡				
Yes	5103 (25.6)	5214 (26.2)	10 317 (25.9)	.20
No	14 832 (74.4)	14 718 (73.8)	29 550 (74.1)	
Hyperlipidemia§				
Yes	5842 (29.3)	5903 (29.6)	11 745 (29.5)	.50
No	14 089 (70.7)	14 026 (70.4)	28 115 (70.5)	
Diabetes‖				
Yes	517 (2.6)	510 (2.6)	1027 (2.6)	.83
No	19 411 (97.4)	19 414 (97.4)	38 825 (97.4)	
Parental history of myocardial infarction before age 60 y				.83
Yes	2321 (13.0)	2312 (12.9)	4633 (12.9)	
No	15583 (87.0)	15 627 (87.1)	31 210 (87.1)	

*Numbers do not always sum to group totals due to missing information for some variables.
†Body mass index was calculated as weight in kilograms divided by the square of height in meters.
‡Hypertension was defined as a self-reported systolic blood pressure ≥140 mm Hg, a diastolic blood pressure ≥90 mm Hg, or physician-diagnosed hypertension.
§Hyperlipidemia was defined as a self-reported total cholesterol ≥240 mg/dL (6.2 mmol/L) or physician-diagnosed high cholesterol.
‖Diabetes was defined by self-report.

Table 2. Relative Risks of Cardiovascular Disease, Cancer, and Total Mortality by Group, Women's Health Study

Outcome	No. of Events		Relative Risk (95% CI)	*P* Value
	Vitamin E (n = 19 937)	Placebo (n = 19 939)		
Major cardiovascular event*	482	517	0.93 (0.82-1.05)	.26
Myocardial infarction	196	195	1.01 (0.82-1.23)	.96
Nonfatal	184	181	1.02 (0.83-1.25)	.87
Fatal	12	14	0.86 (0.40-1.85)	.70
Stroke	241	246	0.98 (0.82-1.17)	.82
Nonfatal	220	222	0.99 (0.82-1.19)	.93
Fatal	21	24	0.88 (0.49-1.57)	.66
Ischemic†	194	197	0.99 (0.81-1.20)	.88
Hemorrhagic†	44	48	0.92 (0.61-1.38)	.68
Cardiovascular death	106	140	0.76 (0.59-0.98)	.03
Total invasive cancer	1437	1428	1.01 (0.94-1.08)	.87
Breast	616	614	1.00 (0.90-1.12)	.95
Lung	107	98	1.09 (0.83-1.44)	.52
Colon	107	107	1.00 (0.77-1.31)	.99
Cancer death	308	275	1.12 (0.95-1.32)	.17
Total mortality	636	615	1.04 (0.93-1.16)	.53

Abbreviation: CI, confidence interval.
*Defined as a composite end point comprising the first of any of these events: nonfatal myocardial infarction, nonfatal stroke, or cardiovascular death.
†Stroke type was not known for 3 women in the vitamin E group and 1 in the placebo group.

100 000 person-years): 1437 in the vitamin E group and 1428 in the placebo group (RR, 1.01; 95% CI, 0.94-1.08; *P*=.87; Table 2). Further analyses that excluded the first 2 or 5 years of follow-up did not change the findings (data not shown). For the main site-specific cancers, there were no significant differences between the vitamin E and placebo groups (breast cancer: RR, 1.00; 95% CI, 0.90-1.12; lung cancer: RR, 1.09; 95% CI, 0.83-1.44; and colon cancer: RR, 1.00; 95% CI, 0.77-1.31). There also was no significant difference in rectal cancer rates (22 vs 33 cases in the vitamin E and placebo groups, respectively). A previous study in a poorly nourished population reported lower rates of stomach cancer among those randomized to a vitamin and mineral cocktail including vitamin E.[14] We did not observe this in our healthy population (14 vs 6 cases of stomach cancer, respectively). Cancer death rates also were not significantly influenced by vitamin E (308 vs 275 cancer deaths, respectively; RR, 1.12; 95% CI, 0.95-1.32; *P*=.17).

Including in situ and invasive cancers in the analysis led to virtually unchanged findings (1626 vs 1615 cases, respectively; RR, 1.01; 95% CI, 0.94-1.08; *P*=.84).

When we examined cumulative incidence rates of invasive cancer by year of follow-up, the curves were almost identical in the vitamin E and placebo groups (FIGURE 3). Additional analyses that censored noncompliant (taking less than two thirds of study medications) follow-up time continued to show a lack of effect of vitamin E on total invasive cancer (RR, 1.01; 95% CI, 0.93-1.09; *P*=.88).

As with major cardiovascular events, random assignment to neither aspirin nor beta carotene modified the effect of vitamin E on the primary or secondary cancer end points. Additionally, there was no significant effect modification by any of the cancer risk factors shown in Table 3.

Total Mortality

By the end of the trial, 636 women in the vitamin E group had died, as had 615 women in the placebo group (RR, 1.04; 95% CI, 0.93-1.16; *P*=.53). The main causes of death, apart from cardiovascular and cancer deaths, were pulmonary diseases (32 vitamin E, 22 placebo); violent deaths, excluding suicide (31 vs 21); and suicide (9 vs 6). None of these causes of deaths was significantly related to vitamin E.

In analysis that censored noncompliant follow-up time, there also was no significant effect of vitamin E (RR, 1.08; 95% CI, 0.90-1.29; *P*=.42).

There was no effect of random assignment to either aspirin or beta carotene on the effect of vitamin E on total mortality. There also was no significant effect of any of the cardiovascular and cancer risk factors in Table 3 on the association of vitamin E with total mortality.

Adverse Effects

We examined whether vitamin E increased adverse effects due to bleeding (gastrointestinal bleeding, hematuria, easy bruising, epistaxis) because of the potential for vitamin E to inhibit platelet function,[52] gastrointestinal symptoms (gastric upset, nausea, diarrhea, constipation), or fatigue. There were no differences between reported adverse effects for any of these variables among women in the 2 groups, apart from a small, but significant, increase in the risk of epistaxis (RR, 1.06; 95% CI, 1.01-1.11; *P*=.02).

COMMENT

The WHS—the largest randomized trial of vitamin E supplementation to date with the longest duration of treatment—adds important information regarding whether vitamin E plays any role in CVD and cancer prevention. In this trial, 600 IU of natural-source vitamin E every other day for 10 years did not provide any statistically significant benefits on the primary end points of major cardiovascular events or cancer in almost 40 000 healthy women. There was, however, a significant 24% reduction in the secondary end point of cardiovascular deaths and a significant 26% reduction in major cardiovascular events among the subgroup of women aged at least 65 years. We observed no significant effect of vitamin E on total mortality.

The finding of no overall effect of vitamin E on CVD is congruent with data

from previous randomized trials. In 2 recent meta-analyses, the pooled RR of CVD with vitamin E treatment was 1.0 (95% CI, 0.94-1.07) in a 2003 analysis[40] and 0.98 (95% CI, 0.94-1.03) in a 2004 analysis.[41] These trials, however, recruited participants at high risk either because of CVD risk factors or preexisting disease. There are few data on populations comparable with the healthy women in the WHS. A recently published trial not included in either meta-analysis, the SU.VI.MAX study,[35] like the WHS, enrolled primarily healthy persons. After 7.5 years, the SU.VI.MAX trial also reported no effect of randomized treatment using a combination of vitamins and minerals, including 30 mg/d of vitamin E, on CVD (RR, 0.97; 95% CI, 0.77-1.20).

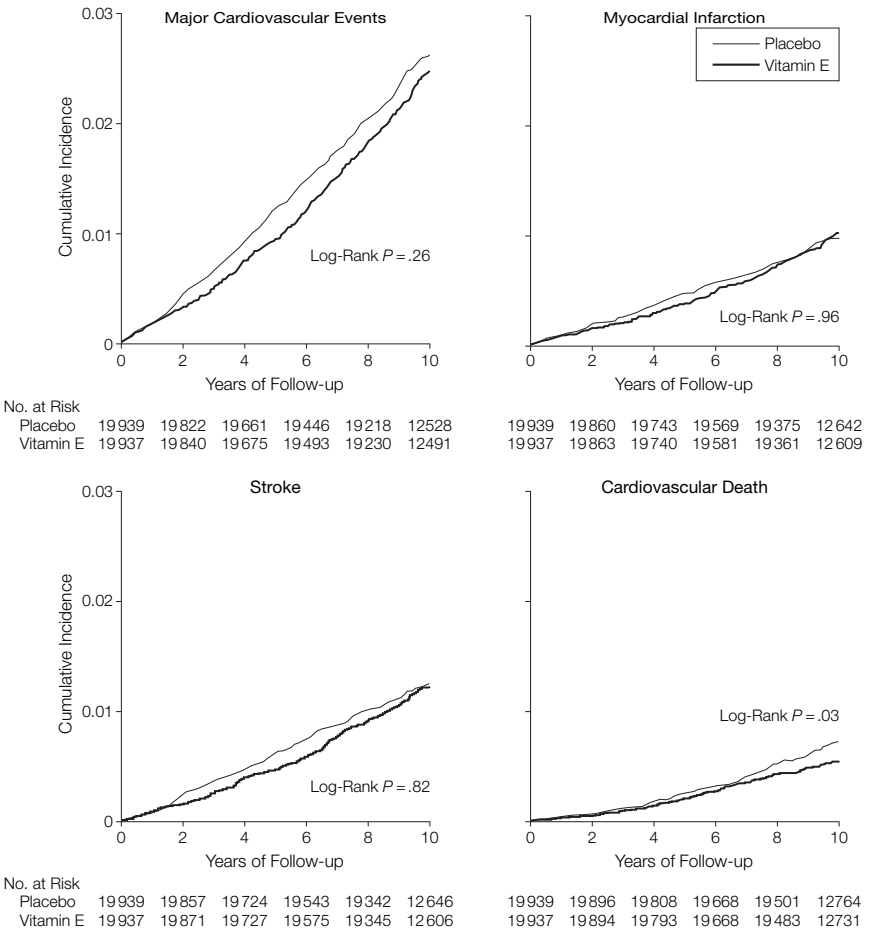
With regard to the individual cardiovascular end points, we found a significant 24% reduction in cardiovascular deaths. This finding differs from the overall data; in the 2003 meta-analysis, the pooled RR for this end point was 1.0 (95% CI, 0.94-1.06)[40] and 1.00 (95% CI, 0.94-1.05) in the 2004 meta-analysis.[41] The addition of the WHS data (106 cardiovascular deaths in the vitamin E group, 140 in the placebo group) to the latter and larger meta-analysis (2683 and 2689 cardiovascular deaths, respectively) should not have an appreciable impact on the pooled RR. In the WHS, the single largest contribution to the reduction in cardiovascular deaths was fewer sudden deaths among women assigned to receive vitamin E. One plausible explanation that we considered was whether omega-3 fatty acids in the treatment capsules may have played a role.[24] This is unlikely, however, because both active and placebo capsules were identically formulated with soybean oil, the only difference being the addition of vitamin E to the active capsules. It is possible that the observed reduction in cardiovascular deaths was due to chance, arising from multiple comparisons.

An interesting finding in subgroup analyses was the observation of a significant 26% reduction in major cardiovascular events, primarily cardiovascular deaths, among women aged at least 65 years. Few previous trials of vitamin E have reported findings by age. The one that did, the HOPE trial, enrolled participants aged at least 55 years with CVD, or diabetes and one other risk factor, and reported no overall effect of vitamin E on CVD and no heterogeneity of results by age.[26] Several large observational studies that noted inverse associations between vitamin E intake and CVD rates did not provide findings by age.[8-10] Existing trials of vitamin E can help clarify this by providing findings regarding any age effects.

A recent trial, HOPE-TOO, noted a possible adverse effect of 400 IU/d of vitamin E on the risk of heart failure.[38] This was not a prespecified end point in the WHS; however, we did collect self-reported information, which did not demonstrate any association between random assignment to vitamin E use and incidence of heart failure. These self-reports are currently being validated against medical records.

In view of the lack of overall benefit of vitamin E on cardiovascular events in the WHS, we considered several factors. First, was the dose of vitamin E sufficient? Previous observational studies have reported significant benefits in women[8] with a median intake of as little as 17 IU/d and 25.2 IU/d in men.[9] The WHS used a far higher dose of 600 IU every other day. Second, the lack of

Figure 2. Cumulative Incidence Rates of Cardiovascular Disease



The composite cardiovascular end point (the first of any of the individual end points) is reported as well as the individual end points of myocardial infarction, stroke, and cardiovascular death.

Table 3. Relative Risks of Cardiovascular Disease and Cancer According to Baseline Characteristics, Women's Health Study

Group	Major Cardiovascular Event*				Total Invasive Cancer			
	No. of Events				No. of Events			
	Vitamin E	Placebo	RR (95% CI)	*P* Value	Vitamin E	Placebo	RR (95% CI)	*P* Value
Age, y†								
45-54	172	152	1.13 (0.91-1.41)	.26	645	673	0.96 (0.86-1.07)	.44
55-64	180	189	0.95 (0.77-1.16)	.61	518	509	1.02 (0.90-1.15)	.79
≥65	130	176	0.74 (0.59-0.93)	.009	274	246	1.12 (0.94-1.33)	.20
Smoking status								
Current	139	145	0.99 (0.78-1.25)	.91	238	228	1.07 (0.89-1.29)	.45
Past or never	341	370	0.92 (0.79-1.06)	.25	1199	1199	1.00 (0.92-1.08)	.95
Alcohol intake								
Never/rarely	252	263	0.96 (0.80-1.14)	.61	643	615	1.04 (0.93-1.16)	.52
At least 1/mo	230	254	0.90 (0.76-1.08)	.26	794	813	0.98 (0.89-1.09)	.75
Multivitamin use								
Yes	201	194	1.02 (0.84-1.25)	.81	577	531	1.07 (0.95-1.21)	.24
No	281	323	0.88 (0.75-1.03)	.10	860	897	0.97 (0.88-1.06)	.48
Body mass index‡								
<25	186	221	0.85 (0.70-1.03)	.09	709	728	0.98 (0.89-1.09)	.75
25-29	164	169	0.96 (0.77-1.19)	.69	429	403	1.05 (0.92-1.21)	.46
≥30	113	114	1.00 (0.77-1.30)	.99	269	272	0.99 (0.84-1.18)	.95
Physical activity, kcal/wk								
<1000	328	357	0.92 (0.79-1.07)	.28	949	967	0.98 (0.89-1.07)	.64
≥1000	147	153	0.95 (0.76-1.20)	.68	473	445	1.08 (0.94-1.22)	.27
Menopausal status and hormone therapy use								
Premenopausal	59	59	1.01 (0.71-1.45)	.94	294	300	1.00 (0.85-1.17)	.96
Uncertain	75	67	1.13 (0.81-1.57)	.46	173	211	0.82 (0.67-1.00)	.06
Postmenopausal, current hormone therapy use	145	154	0.95 (0.76-1.20)	.68	507	461	1.11 (0.98-1.26)	.11
Postmenopausal, never or past hormone therapy use	201	234	0.84 (0.70-1.02)	.08	460	453	1.00 (0.88-1.14)	.97
Hypertension§								
Yes	242	263	0.93 (0.78-1.11)	.41	408	424	0.97 (0.85-1.11)	.68
No	239	254	0.94 (0.78-1.12)	.46	1028	1003	1.02 (0.94-1.11)	.63
Hyperlipidemia‖								
Yes	189	221	0.86 (0.71-1.05)	.14	493	460	1.09 (0.96-1.23)	.20
No	293	296	0.98 (0.83-1.15)	.82	944	968	.97 (0.89-1.06)	.50
Diabetes¶								
Yes	62	58	1.05 (0.73-1.50)	.79	48	35	1.39 (0.90-2.15)	.14
No	420	458	0.92 (0.80-1.04)	.19	1388	1393	1.00 (0.93-1.07)	.93
Randomized to receive aspirin								
Yes	232	245	0.95 (0.79-1.13)	.55	716	722	0.99 (0.89-1.10)	.86
No	250	272	0.92 (0.77-1.09)	.32	721	706	1.02 (0.92-1.13)	.67
Parental history of MI before age 60 y								
Yes	73	62	1.16 (0.82-1.62)	.40				
No	352	391	0.90 (0.78-1.04)	.16				
10-Year risk of CHD, %#								
<5.0	179	148	1.19 (0.96-1.48)	.12				
5.0-9.9	98	95	1.10 (0.83-1.46)	.50				
≥10.0	53	68	0.81 (0.56-1.16)	.24				

Abbreviations: CHD, coronary heart disease; CI, confidence interval; MI, myocardial infarction; RR, relative risk.
*Defined as a composite end point comprising the first of any of these events: nonfatal MI, nonfatal stroke, or cardiovascular death.
†*P* value for interaction <.05 for major cardiovascular event. No other interactions are significant.
‡Calculated as weight in kilograms divided by the square of height in meters.
§Hypertension was defined as a self-reported systolic blood pressure ≥140 mm Hg, a diastolic blood pressure ≥90 mm Hg, or physician-diagnosed hypertension.
‖Hyperlipidemia was defined as a self-reported total cholesterol ≥240 mg/dL (6.2 mmol/L) or physician-diagnosed high cholesterol.
¶Diabetes defined by self-report.
#Calculated using the Framingham risk score among 28 345 women who provided a blood sample at baseline.

effect was unlikely to be due to insufficient treatment duration, since this averaged 10 years in the WHS, representing the longest duration of any vitamin E trial. Third, we considered whether the source of vitamin E used, a natural source, influenced the findings. This was unlikely; 2 previous trials of secondary prevention that reported a benefit of vitamin E also used a natural source.[21,27] On the other hand, a previous trial of secondary prevention, which included high-risk primary prevention patients, used natural-source vitamin E and found no benefits on CVD,[26] as did another secondary prevention trial testing natural-source vitamin E combined with other vitamins and minerals.[31]

Fourth, declining compliance over time in the WHS may have diluted the findings. However, in sensitivity analyses in which follow-up time was censored among women taking less than two thirds of their study pills, the finding for cardiovascular events was little changed. Additionally, accounting for outside use of vitamin E also did not make a difference.

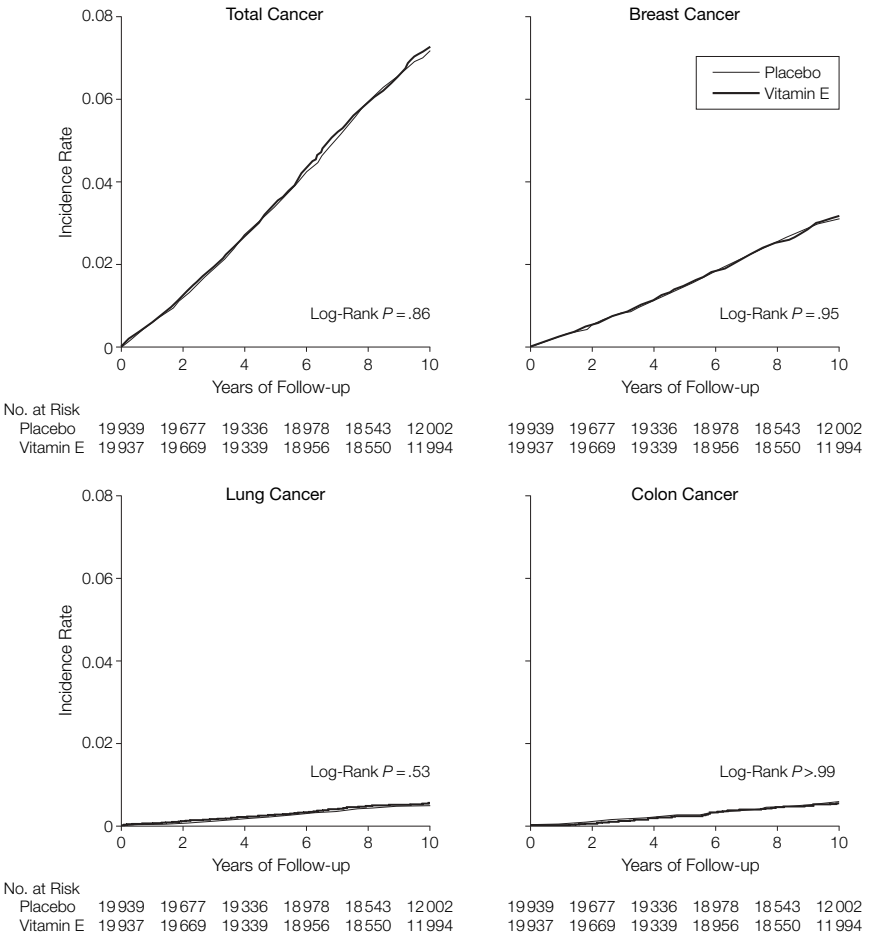
Fifth, the hypothesis has been raised that antioxidants may adversely interact with simvastatin and niacin treatment.[31] We did not systematically collect information on lipid-modifying therapy, but we did so for hyperlipidemia. Among women who remained normolipemic throughout the trial and who were unlikely to have taken lipid-modifying drugs, we observed no significant effect of vitamin E on major cardiovascular events, providing little support for an influence of lipid therapy on the WHS findings.

Finally, the possibility exists that γ-tocopherol, rather than vitamin E (or α-tocopherol), may be the relevant compound for CVD prevention.[53] γ-Tocopherol appears to have similar or greater efficacy than α-tocopherol at inhibiting lipid peroxidation under oxyradical systems and much more potency using nitration systems.[53]

With regard to the cancer end points, there are few data from randomized trials of vitamin E.[43,44,54] The ATBC trial, conducted among men, observed a lower incidence of prostate cancer among men assigned to receive 50 mg/d of vitamin E, but no effect on lung or colon cancers.[16,17] In the HOPE-TOO trial, there was no significant effect of 400 IU/d of vitamin E on cancer incidence or deaths, as in the WHS.[38] There was a lower incidence of lung cancer with vitamin E in HOPE-TOO, not reaching the predefined level of statistical significance. We did not observe any effect of vitamin E on lung cancer in the WHS. The SU.VI.MAX study reported significantly lower cancer rates among men, but not women, randomized to a combination of vitamins and minerals (including 30 mg/d of vitamin E).[35] Among poorly nourished persons randomized to a vitamin and mineral cocktail (including 30 mg/d of vitamin E), lower rates of stomach cancer occurred.[14] This was not seen in the WHS, but the number of stomach cancers was small. Taken as a whole, the available data do not provide strong evidence for a role of vitamin E in cancer prevention, particularly in well-nourished persons.

A recent meta-analysis raised concern for increased mortality with vitamin E, especially in doses of 400 IU/d or greater.[42] In the WHS, using 600 IU every other day, there was no significant effect of vitamin E on total mortality. There was no excess of cardiovascular (and, indeed, fewer such deaths) or cancer deaths, the main causes of mortality, in the vitamin E group. For the other main causes of

Figure 3. Cumulative Incidence Rates of Cancer



death, there were more deaths (but not statistically significant) from pulmonary disease, violent deaths, and suicides in the vitamin E group.

Vitamin E was well tolerated in the WHS with no significant differences in adverse effects between groups, except for epistaxis. This is likely to be a chance finding because there were no differences in other adverse effects from bleeding. Noteworthy was the observation of no increase in hemorrhagic strokes with vitamin E, in contrast to the ATBC trial with an excess of deaths from such strokes.[16]

CONCLUSIONS

In conclusion, the WHS does not support recommending vitamin E supplementation for CVD or cancer prevention among healthy women. This large trial supports current guidelines stating that use of antioxidant vitamins is not justified for CVD risk reduction.[55,56] The WHS finding of a decreased cardiovascular death rate with vitamin E, as well as decreased major cardiovascular events among women aged at least 65 years, differs from the totality of evidence and should be explored further. The WHS findings should be viewed in the context of the available randomized evidence, as well as data that should be available over the next several years from ongoing trials, including the Physicians' Health Study, which will provide data on primary prevention in men.[57] At present, in the primary prevention of CVD and cancer, therapeutic lifestyle changes including a healthy diet and control of major risk factors remain important clinical and public health strategies.

Author Affiliations: Divisions of Preventive Medicine (Drs Lee, Cook, Gaziano, Ridker, Manson, and Buring and Mr Gordon), Cardiovascular Medicine (Drs Gaziano and Ridker), and Aging (Drs Gaziano and Buring), Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Mass; Department of Epidemiology, Harvard School of Public Health, Boston (Drs Lee, Cook, Ridker, Manson, and Buring); Veterans Affairs Boston Healthcare System (Dr Gaziano); Department of Ambulatory Care and Prevention, Harvard Medical School, Boston (Dr Buring); and the Departments of Medicine and Epidemiology and Public Health, University of Miami School of Medicine, Miami, Fla, and Department of Biomedical Science, Center of Excellence, Florida Atlantic University, Boca Raton (Dr Hennekens).

Author Contributions: Dr Lee had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Lee, Cook, Gaziano, Gordon, Ridker, Manson, Hennekens, Buring.
Acquisition of data: Lee, Gaziano, Gordon, Hennekens, Buring.
Analysis and interpretation of data: Lee, Cook, Gaziano, Ridker, Manson, Hennekens, Buring.
Drafting of the manuscript: Lee.
Critical revision of the manuscript for important intellectual content: Lee, Cook, Gaziano, Gordon, Ridker, Manson, Hennekens, Buring.
Statistical analysis: Cook.
Obtained funding: Hennekens, Buring.
Administrative, technical, or material support: Lee, Gaziano, Gordon, Ridker, Manson, Hennekens, Buring.
Study supervision: Lee, Gaziano, Gordon, Ridker, Manson, Hennekens, Buring.
Financial Disclosures: Dr Cook has served as a consultant to Bayer. Dr Gaziano has served as a consultant to and received grant support from Bayer, DSM Nutritional Products Inc, and McNeil. Dr Ridker has received grant support from Bayer. Dr Hennekens has served as a consultant to Bayer and McNeil and received grant support from Bayer.
Funding/Support: This study was supported by grants HL-43851 and CA-47988 from the National Heart, Lung, and Blood Institute and the National Cancer Institute. Vitamin E and vitamin E placebo were provided by the Natural Source Vitamin E Association. Aspirin and aspirin placebo (reported on pp 47-55) were provided by Bayer Healthcare.
Role of the Sponsor: Neither the Natural Source Vitamin E Association nor Bayer Healthcare provided any input into the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.
Data and Safety Monitoring Board: Lawrence Cohen, Rory Collins, Theodore Colton, David DeMets, I. Craig Henderson, Andrea La Croix, Ross Prentice, and Nanette Wenger (chair) and Mary Frances Cotch, Frederick Ferris, Lawrence Friedman, Peter Greenwald, Natalie Kurinij, Marjorie Perloff, Eleanor Schron, and Alan Zonderman (ex-officio members).
Acknowledgment: We are indebted to the 39 876 participants in the Women's Health Study for their dedicated and conscientious collaboration; to the entire staff of the Women's Health Study, under the leadership of David Gordon, Maria Andrade, Susan Burt, Mary Breen, Marilyn Chown, Lisa Fields-Johnson, Georgina Friedenberg, Inge Judge, Jean MacFadyen, Geneva McNair, Laura Pestana, Philomena Quinn, Claire Ridge, Harriet Samuelson, Fred Schwerin and Marty Van Denburgh; to Christine Albert, Michelle Albert, Gavin Blake, Claudia Chae, Wendy Chen, Richard Doll, Carlos Kase, Tobias Kurth, Richard Peto, Aruna Pradhan, Kathryn Rexrode, Bernard Rosner, Jacqueline Suk, and Shumin Zhang for their assistance in the design and conduct of the trial; and especially to James Taylor for chairing the Endpoints Committee.

REFERENCES

1. Halliwell B. Free radicals, antioxidants, and human disease: curiosity, cause, or consequence? *Lancet*. 1994;344:721-724.
2. Steinberg D, Parthasarathy S, Carew TE, Khoo JC, Witztum JL. Beyond cholesterol:. modifications of low-density lipoprotein that increase its atherogenicity. *N Engl J Med*. 1989;320:915-924.
3. Witztum JL. The oxidation hypothesis of atherosclerosis. *Lancet*. 1994;344:793-795.
4. Cerutti PA. Oxy-radicals and cancer. *Lancet*. 1994; 344:862-863.
5. Meydani M. Vitamin E. *Lancet*. 1995;345:170-175.
6. Upston JM, Terentis AC, Stocker R. Tocopherol-mediated peroxidation of lipoproteins: implications for vitamin E as a potential antiatherogenic supplement. *FASEB J*. 1999;13:977-994.
7. Gey KF, Puska P, Jordan P, Moser UK. Inverse correlation between plasma vitamin E and mortality from ischemic heart disease in cross-cultural epidemiology. *Am J Clin Nutr*. 1991;53(1 suppl):326S-334S.
8. Stampfer MJ, Hennekens CH, Manson JE, Colditz GA, Rosner B, Willett WC. Vitamin E consumption and the risk of coronary disease in women. *N Engl J Med*. 1993;328:1444-1449.
9. Rimm EB, Stampfer MJ, Ascherio A, Giovannucci E, Colditz GA, Willett WC. Vitamin E consumption and the risk of coronary heart disease in men. *N Engl J Med*. 1993;328:1450-1456.
10. Kushi LH, Folsom AR, Prineas RJ, Mink PJ, Wu Y, Bostick RM. Dietary antioxidant vitamins and death from coronary heart disease in postmenopausal women. *N Engl J Med*. 1996;334:1156-1162.
11. Mehta J. Intake of antioxidants among American cardiologists. *Am J Cardiol*. 1997;79:1558-1560.
12. Patterson RE, White E, Kristal AR, Neuhouser ML, Potter JD. Vitamin supplements and cancer risk: the epidemiologic evidence. *Cancer Causes Control*. 1997; 8:786-802.
13. Hennekens CH, Buring JE. *Epidemiology in Medicine*. Boston, Mass: Little Brown & Co; 1987.
14. Blot WJ, Li JY, Taylor PR, et al. Nutrition intervention trials in Linxian, China: supplementation with specific vitamin/mineral combinations, cancer incidence, and disease-specific mortality in the general population. *J Natl Cancer Inst*. 1993;85: 1483-1492.
15. Li JY, Taylor PR, Li B, et al. Nutrition intervention trials in Linxian, China: multiple vitamin/mineral supplementation, cancer incidence, and disease-specific mortality among adults with esophageal dysplasia. *J Natl Cancer Inst*. 1993;85:1492-1498.
16. The Alpha-Tocopherol, Beta Carotene Cancer Prevention Study Group. The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers. *N Engl J Med*. 1994;330: 1029-1035.
17. Virtamo J, Pietinen P, Huttunen JK, et al. Incidence of cancer and mortality following alpha-tocopherol and beta-carotene supplementation: a postintervention follow-up. *JAMA*. 2003;290:476-485.
18. Rapola JM, Virtamo J, Ripatti S, et al. Randomised trial of alpha-tocopherol and beta-carotene supplements on incidence of major coronary events in men with previous myocardial infarction. *Lancet*. 1997;349:1715-1720.
19. Virtamo J, Rapola JM, Ripatti S, et al. Effect of vitamin E and beta carotene on the incidence of primary nonfatal myocardial infarction and fatal coronary heart disease. *Arch Intern Med*. 1998;158:668-675.
20. Greenberg ER, Baron JA, Tosteson TD, et al; Polyp Prevention Study Group. A clinical trial of antioxidant vitamins to prevent colorectal adenoma. *N Engl J Med*. 1994;331:141-147.
21. Stephens NG, Parsons A, Schofield PM, Kelly F, Cheeseman K, Mitchinson MJ. Randomised controlled trial of vitamin E in patients with coronary disease: Cambridge Heart Antioxidant Study (CHAOS). *Lancet*. 1996;347:781-786.
22. Sano M, Ernesto C, Thomas RG, et al; The Alzheimer's Disease Cooperative Study. A controlled trial of selegiline, alpha-tocopherol, or both as treatment for Alzheimer's disease. *N Engl J Med*. 1997;336:1216-1222.
23. Parkinson Study Group. Mortality in DATATOP: a multicenter trial in early Parkinson's disease. *Ann Neurol*. 1998;43:318-325.
24. Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto miocardico. Dietary supplementation with n-3 polyunsaturated fatty acids and vita-

min E after myocardial infarction: results of the GISSI-Prevenzione trial. *Lancet*. 1999;354:447-455.
25. Girodon F, Galan P, Monget AL, et al. Impact of trace elements and vitamin supplementation on immunity and infections in institutionalized elderly patients: a randomized controlled trial. *Arch Intern Med*. 1999;159:748-754.
26. Yusuf S, Dagenais G, Pogue J, Bosch J, Sleight P; The Heart Outcomes Prevention Evaluation Study Investigators. Vitamin E supplementation and cardiovascular events in high-risk patients. *N Engl J Med*. 2000;342:154-160.
27. Boaz M, Smetana S, Weinstein T, et al. Secondary prevention with antioxidants of cardiovascular disease in endstage renal disease (SPACE): randomised placebo-controlled trial. *Lancet*. 2000;356:1213-1218.
28. Collaborative Group of the Primary Prevention Project. Low-dose aspirin and vitamin E in people at cardiovascular risk: a randomised trial in general practice. *Lancet*. 2001;357:89-95.
29. You WC, Chang YS, Heinrich J, et al. An intervention trial to inhibit the progression of precancerous gastric lesions: compliance, serum micronutrients and S-allyl cysteine levels, and toxicity. *Eur J Cancer Prev*. 2001;10:257-263.
30. A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E and beta carotene for age-related cataract and vision loss: AREDS report no. 9. *Arch Ophthalmol*. 2001;119:1439-1452.
31. Brown BG, Zhao XQ, Chait A, et al. Simvastatin and niacin, antioxidant vitamins, or the combination for the prevention of coronary disease. *N Engl J Med*. 2001;345:1583-1592.
32. Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of antioxidant vitamin supplementation in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet*. 2002;360:23-33.
33. Chylack LT Jr, Brown NP, Bron A, et al. The Roche European American Cataract Trial (REACT): a randomized clinical trial to investigate the efficacy of an oral antioxidant micronutrient mixture to slow progression of age-related cataract. *Ophthalmic Epidemiol*. 2002;9:49-80.
34. Waters DD, Alderman EL, Hsia J, et al. Effects of hormone replacement therapy and antioxidant vitamin supplements on coronary atherosclerosis in postmenopausal women: a randomized controlled trial. *JAMA*. 2002;288:2432-2440.
35. Hercberg S, Galan P, Preziosi P, et al. The SU.VI-.MAX Study: a randomized, placebo-controlled trial of the health effects of antioxidant vitamins and minerals. *Arch Intern Med*. 2004;164:2335-2342.
36. Zureik M, Galan P, Bertrais S, et al. Effects of long-term daily low-dose supplementation with antioxidant vitamins and minerals on structure and function of large arteries. *Arterioscler Thromb Vasc Biol*. 2004;24:1485-1491.
37. McNeil JJ, Robman L, Tikellis G, Sinclair MI, McCarty CA, Taylor HR. Vitamin E supplementation and cataract: randomized controlled trial. *Ophthalmology*. 2004;111:75-84.
38. Lonn E, Bosch J, Yusuf S, et al. Effects of long-term vitamin E supplementation on cardiovascular events and cancer: a randomized controlled trial. *JAMA*. 2005;293:1338-1347.
39. Manson JE, Bassuk SS, Stampfer MJ. Does vitamin E supplementation prevent cardiovascular events? *J Womens Health (Larchmt)*. 2003;12:123-136.
40. Vivekananthan DP, Penn MS, Sapp SK, Hsu A, Topol EJ. Use of antioxidant vitamins for the prevention of cardiovascular disease: meta-analysis of randomised trials. *Lancet*. 2003;361:2017-2023.
41. Eidelman RS, Hollar D, Hebert PR, Lamas GA, Hennekens CH. Randomized trials of vitamin E in the treatment and prevention of cardiovascular disease. *Arch Intern Med*. 2004;164:1552-1556.
42. Miller ER III, Pastor-Barriuso R, Dalal D, Riemersma RA, Appel LJ, Guallar E. Meta-analysis: high-dosage vitamin E supplementation may increase all-cause mortality. *Ann Intern Med*. 2005;142:37-46.
43. Ritenbaugh C, Streit K, Helfand M. Summary of evidence from randomized controlled trials: routine vitamin supplementation to prevent cancer. Rockville, Md: Agency for Healthcare Research and Quality. Available at: http://www.ahrq.gov/clinic/3rduspstf/vitamins/vitasum.htm. Accessed May 18, 2005.
44. Shetty P, Atkins D. Update of the evidence from randomized controlled trials, 1999-2002: routine vitamin supplementation to prevent cancer. Rockville, Md: Agency for Healthcare Research and Quality. Available at: http://www.ahrq.gov/clinic/3rduspstf/vitamins/vitupdate.htm. Accessed May 18, 2005.
45. Steinberg D. Clinical trials of antioxidants in atherosclerosis: are we doing the right thing? *Lancet*. 1995;346:36-38.
46. Rexrode KM, Lee IM, Cook NR, Hennekens CH, Buring JE. Baseline characteristics of participants in the Women's Health Study. *J Womens Health Gend Based Med*. 2000;9:19-27.
47. Ridker PM, Cook NR, Lee IM, et al. A randomized trial of low-dose aspirin in the primary prevention of cardiovascular disease in women. *N Engl J Med*. 2005;352:1293-1304.
48. Lee IM, Cook NR, Manson JE, Buring JE, Hennekens CH. Beta-carotene supplementation and incidence of cancer and cardiovascular disease: the Women's Health Study. *J Natl Cancer Inst*. 1999;91:2102-2106.
49. Atiya M, Kurth T, Berger K, Buring JE, Kase CS. Interobserver agreement in the classification of stroke in the Women's Health Study. *Stroke*. 2003;34:565-567.
50. Halliwell B. The antioxidant paradox. *Lancet*. 2000;355:1179-1180.
51. Ogita H, Liao J. Endothelial function and oxidative stress. *Endothelium*. 2004;11:123-132.
52. Steiner M. Vitamin E, a modifier of platelet function: rationale and use in cardiovascular and cerebrovascular disease. *Nutr Rev*. 1999;57:306-309.
53. Hensley K, Benaksas EJ, Bolli R, et al. New perspectives on vitamin E: gamma-tocopherol and carboxyelthylhydroxychroman metabolites in biology and medicine. *Free Radic Biol Med*. 2004;36:1-15.
54. Bairati I, Meyer F, Gelinas M, et al. A randomized trial of antioxidant vitamins to prevent second primary cancers in head and neck cancer patients. *J Natl Cancer Inst*. 2005;97:481-488.
55. Mosca L, Appel LJ, Benjamin EJ, et al. Evidence-based guidelines for cardiovascular disease prevention in women. *Circulation*. 2004;109:672-693.
56. Kris-Etherton PM, Lichtenstein AH, Howard BV, Steinberg D, Witztum JL. Antioxidant vitamin supplements and cardiovascular disease. *Circulation*. 2004;110:637-641.
57. Christen WG, Gaziano JM, Hennekens CH. Design of Physicians' Health Study II: a randomized trial of beta-carotene, vitamins E and C, and multivitamins, in prevention of cancer, cardiovascular disease, and eye disease, and review of results of completed trials. *Ann Epidemiol*. 2000;10:125-134.

Table 1. STARD checklist for the reporting of studies of diagnostic accuracy.

Section and Topic	Item #		On page #
TITLE/ABSTRACT/ KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity').	
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	
METHODS		Describe	
Participants	3	The study population: The inclusion and exclusion criteria, setting and locations where the data were collected.	
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in items 3 and 4? If not, specify how participants were further selected.	
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	
Test methods	7	The reference standard and its rationale.	
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	
	9	Definition of and rationale for the units, cutoffs and/or categories of the results of the index tests and the reference standard.	
	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard.	
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	
	13	Methods for calculating test reproducibility, if done.	
RESULTS		Report	
Participants	14	When study was done, including beginning and ending dates of recruitment.	
	15	Clinical and demographic characteristics of the study population (e.g. age, sex, spectrum of presenting symptoms, comorbidity, current treatments, recruitment centers).	
	16	The number of participants satisfying the criteria for inclusion that did or did not undergo the index tests and/or the reference standard; describe why participants failed to receive either test (a flow diagram is strongly recommended).	
Test results	17	Time interval from the index tests to the reference standard, and any treatment administered between.	
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	
	20	Any adverse events from performing the index tests or the reference standard.	
Estimates	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).	
	22	How indeterminate results, missing responses and outliers of the index tests were handled.	
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	
	24	Estimates of test reproducibility, if done.	
DISCUSSION	25	Discuss the clinical applicability of the study findings.	

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses

Continued on next page

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

MOOSE checklist for reporting and appraising meta-analyses of observational studies

Reporting of background should include:

- Problem definition
- Hypothesis statement
- Description of study outcome(s)
- Type of exposure or intervention used
- Type of study design used
- Study population

Reporting of search strategy should include:

- Qualifications of searchers (eg, librarians and investigators)
- Search strategy, including time period inclusion in the synthesis and keywords
- Effort to include all available studies, including contact with authors
- Database and registries searched
- Search software used, name of version, including special features used (eg, explosion)
- Use of hand searching (eg, reference lists of obtained articles)
- List of citations located and those excluded, including justification
- Method of addressing articles published in languages other than English
- Method of handling abstracts and unpublished studies
- Description of any contact with authors

Reporting of methods should include:

- Description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested
- Rationale for the selection and coding of data (eg, sound, clinical principles or convenience)
- Documentation of how data were classified and coded (eg, multiple raters, blinding, and interrater reliability)
- Assessment of confounding (eg, comparability of cases and controls in studies where appropriate)
- Assessment of study quality, including blinding of quality assessors; stratification or regression on possible predictors of study results
- Assessment of heterogeneity
- Description of statistical methods (eg, complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient detail to be replicated
- Provision of appropriate tables and graphics

Reporting of results should include:

- Graphic summarizing individual study estimates and overall estimate
- Table given descriptive information for watch study included
- Results of sensitivity testing (eg, subgroup analysis)
- Indication of statistical uncertainty of findings

Reporting of discussion should include:

- Quantitative assessment if bias (eg, publication bias)
- Justification for exclusion (eg, exclusion of non-English-language citations)
- Assessment of quality of included studies

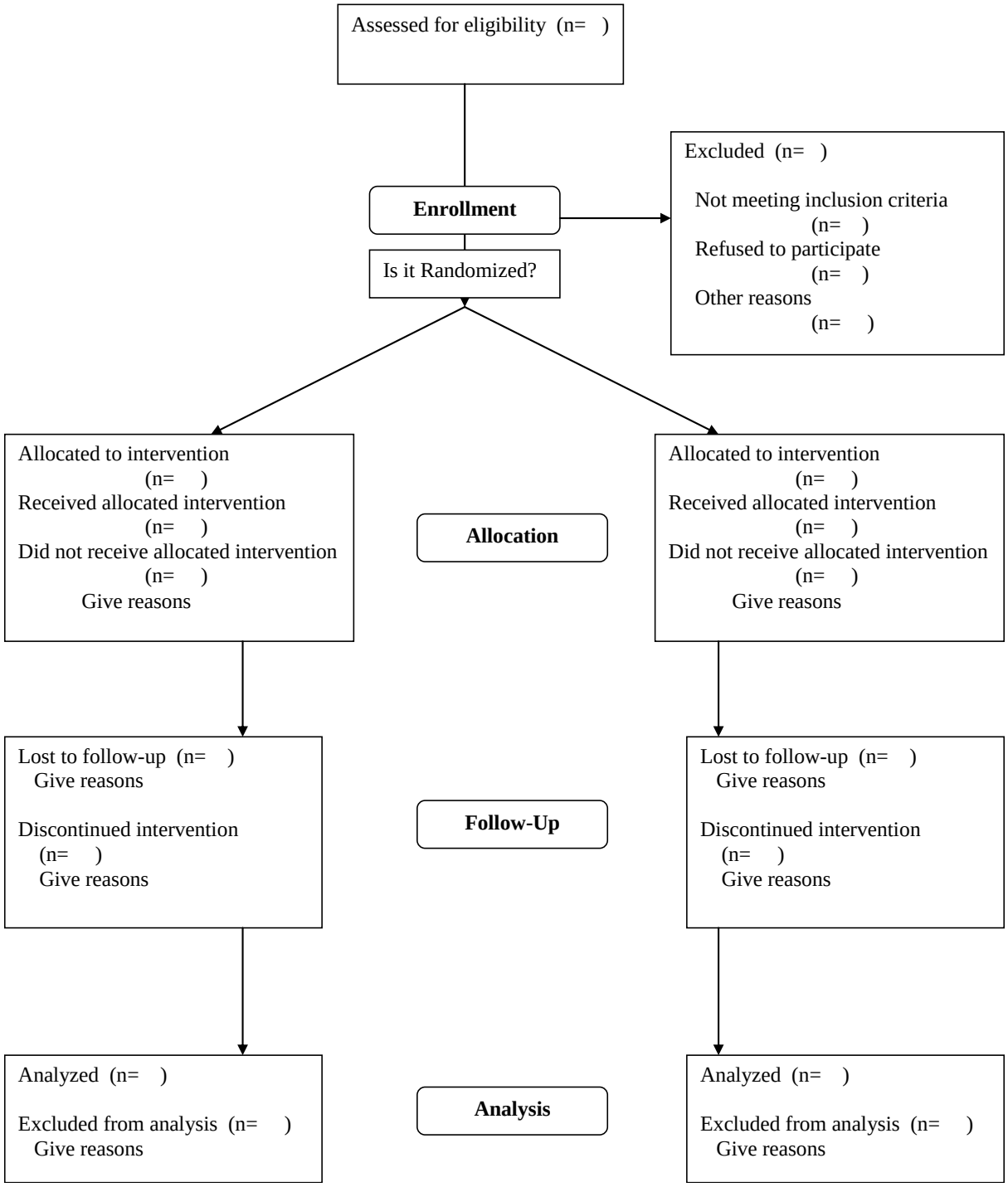
Reporting of conclusions should include:

- Consideration of alternative explanations for observed results
- Generalization of conclusions (ie, appropriate for the data presented and within the domain of the literature review)
- Guidelines for future research
- Disclosure of funding source

CONSORT STATEMENT 2001 - Checklist
Items to include when reporting a randomized trial

PAPER SECTION And topic	Item	Description	Reported on Page #
TITLE & ABSTRACT	1	How participants were allocated to interventions (*e.g.*, "random allocation", "randomized", or "randomly assigned").	
INTRODUCTION Background	2	Scientific background and explanation of rationale.	
METHODS Participants	3	Eligibility criteria for participants and the settings and locations where the data were collected.	
Interventions	4	Precise details of the interventions intended for each group and how and when they were actually administered.	
Objectives	5	Specific objectives and hypotheses.	
Outcomes	6	Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (*e.g.*, multiple observations, training of assessors).	
Sample size	7	How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules.	
Randomization -- Sequence generation	8	Method used to generate the random allocation sequence, including details of any restrictions (*e.g.*, blocking, stratification)	
Randomization -- Allocation concealment	9	Method used to implement the random allocation sequence (*e.g.*, numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned.	
Randomization -- Implementation	10	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups.	
Blinding (masking)	11	Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. When relevant, how the success of blinding was evaluated.	
Statistical methods	12	Statistical methods used to compare groups for primary outcome(s); Methods for additional analyses, such as subgroup analyses and adjusted analyses.	
RESULTS Participant flow	13	Flow of participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome. Describe protocol deviations from study as planned, together with reasons.	
Recruitment	14	Dates defining the periods of recruitment and follow-up.	
Baseline data	15	Baseline demographic and clinical characteristics of each group.	
Numbers analyzed	16	Number of participants (denominator) in each group included in each analysis and whether the analysis was by "intention-to-treat". State the results in absolute numbers when feasible (*e.g.*, 10/20, not 50%).	
Outcomes and estimation	17	For each primary and secondary outcome, a summary of results for each group, and the estimated effect size and its precision (*e.g.*, 95% confidence interval).	
Ancillary analyses	18	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those pre-specified and those exploratory.	
Adverse events	19	All important adverse events or side effects in each intervention group.	
DISCUSSION Interpretation	20	Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision and the dangers associated with multiplicity of analyses and outcomes.	
Generalizability	21	Generalizability (external validity) of the trial findings.	
Overall evidence	22	General interpretation of the results in the context of current evidence.	

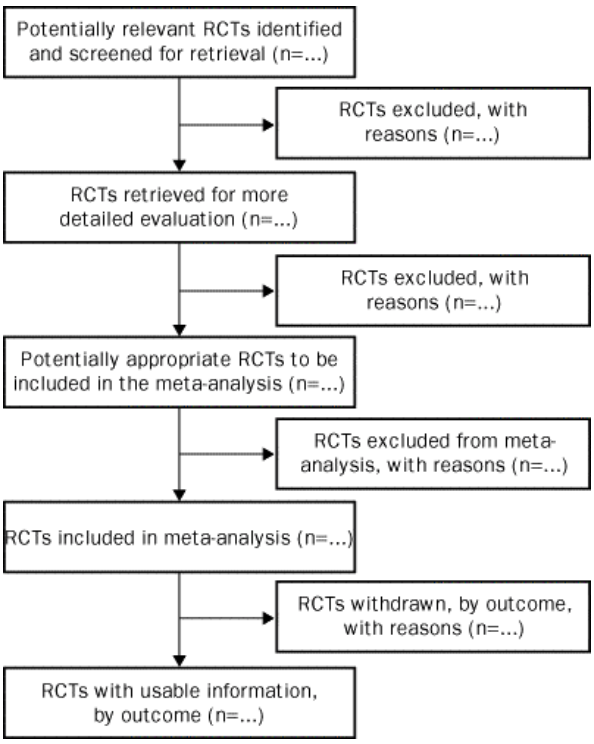
Figure 2. The Consort E-Flowchart



Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement checklist

Heading	Subheading	Descriptor	Reported? (Y/N)	Page number
Title		Identify the report as a meta-analysis [or systematic review] of RCTs[26]		
Abstract		Use a structured format[27]		
		Describe		
	Objectives	The clinical question explicitly		
	Data sources	The databases (ie, list) and other information sources		
	Review methods	The selection criteria (ie, population, intervention, outcome, and study design); methods for validity assessment, data abstraction, and study characteristics, and quantitative data synthesis in sufficient detail to permit replication		
	Results	Characteristics of the RCTs included and excluded; qualitative and quantitative findings (ie, point estimates and confidence intervals); and subgroup analyses		
	Conclusion	The main results		
		Describe		
Introduction		The explicit clinical problem, biological rationale for the intervention, and rationale for review		
Methods	Searching	The information sources, in detail[28] (eg, databases, registers, personal files, expert informants, agencies, hand-searching), and any restrictions (years considered, publication status,[29] language of publication[30,31])		
	Selection	The inclusion and exclusion criteria (defining population, intervention, principal outcomes, and study design[32]		
	Validity assessment	The criteria and process used (eg, masked conditions, quality assessment, and their findings[33–36])		
	Data abstraction	The process or processes used (eg, completed independently, in duplicate)[35,36]		
	Study characteristics	The type of study design, participants' characteristics, details of intervention, outcome definitions, &c,[37] and how clinical heterogeneity was assessed		
	Quantitative data synthesis	The principal measures of effect (eg, relative risk), method of combining results (statistical testing and confidence intervals), handling of missing data; how statistical heterogeneity was assessed;[38] a rationale for any a-priori sensitivity and subgroup analyses; and any assessment of publication bias[39]		
Results	Trial flow	Provide a meta-analysis profile summarising trial flow (see figure)		
	Study characteristics	Present descriptive data for each trial (eg, age, sample size, intervention, dose, duration, follow-up period)		
	Quantitative data synthesis	Report agreement on the selection and validity assessment; present simple summary results (for each treatment group in each trial, for each primary outcome); present data needed to calculate effect sizes and confidence intervals in intention-to-treat analyses (eg 2×2 tables of counts, means and SDs, proportions)		
Discussion		Summarise key findings; discuss clinical inferences based on internal and external validity; interpret the results in light of the totality of available evidence; describe potential biases in the review process (eg, publication bias); and suggest a future research agenda		

Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement flow diagram



*_The Lancet_ is happy for readers to make copies of the checklist and flow diagram. Permission need not be obtained from the journal for reproduction of these items.

NOTES