THE EPIDEMIOLOGY OF ALIMENTARY DISEASES

The Epidemiology of Alimentary Diseases

by

JOHN M. DUGGAN
University of Newcastle, Australia

and

ANNE E. DUGGAN
John Hunter Hospital,
Newcastle, Australia

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-3839-9 (HB)
ISBN-13 978-1-4020-3839-6 (HB)
ISBN-10 1-4020-3840-2 (e-book)
ISBN-13 978-1-4020-3840-2 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springeronline.com

Printed on acid-free paper

Printed in the Netherlands.

Dedication

This book is dedicated to
M.C.D who ultimately made all
this possible.

J.M.D

A.E.D

Contents

Contributing Authors

J.M. Duggan AM
M.D., B.S.(Hons), FRCP, FRACP, FRACMA
Conjoint Professor, School of Medical Practice and Population Health, Faculty of Health, University of Newcastle, Australia.

A.E. Duggan
B.A.(Hons), Dip. Ed., B. Med., M.H.P., Ph.D., FRACP
Conjoint Associate Professor, School of Medical Practice and Population Health, Faculty of Health, University of Newcastle, Australia.

Preface

In offering this book to what we hope will be interested readers, we have several aspirations. We have aspired to present to students and clinicians a rather narrow view of epidemiology concentrating on the causal factors and setting of the more usual gastroenterological problems and giving greater space to conditions of importance for which major knowledge of causation and course is available. Part of the rationale is the belief that modern medicine lays excessive emphasis on therapy with increasingly expensive, and in many cases, dangerous drugs and too little emphasis on the causes and avoidance of disease. We are of the view that traditional views handed down through generations of clinicians need scrutiny worthy of 21st century medicine whose currency includes topics like nanomoles, megabytes and logistic regression. We hope that clinicians will see that there is often a practical application to the findings of epidemiological exploration and that what passes for canonical knowledge is so often unsubstantiated myth and are fully aware of the reluctance of organized medicine to reject old paradigms in favor of the new, matched by an often uncritical enthusiasm for new therapies.

Our researches have increased our belief in the major role of social factors especially diet, both in quantity and quality in many disorders and that clinicians have a responsibility to provide appropriate advice to policy makers as well as patients.

Aware of effect of exposure of undergraduates to the social determinants of disease on their post graduate modes of thinking, we hope that some will even seek to extend their horizons of medicine.

It has been decided that a critical examination of the results of intervention, a part of the landscape of clinical epidemiology is not within

our mandate although it clearly is a field already receiving the deep ploughing it warrants.

Thomas Chalmers, later President of the American Gastroenterological association, in an editorial in Gastroenterology as far back as 1964 made a plea for greater integration of clinical and scientific medicine, quoting Franklin White from a century ago "the new therapy has reduced mortality sharply, but the number of funerals seems to have remained the same". His conclusion was that if clinicians were willing to test hypotheses properly, the funeral rate would drop with the mortality rate (1).

To the public health worker and health policy bureaucrats we offer what we hope will be some evidence to catalyze what must be their frustrating task of trying to bring their expertise into mainstream medicine.

In showing a little of the historical background of some present concepts, maybe we have helped develop interest in what may be conceived as an uninteresting aspect of medicine. If a John Snow should emerge from one of our readers then our efforts will have been fully justified.

JMD

AED

Reference

1. Chalmers TC. Editorial. Mortality rate versus funeral rate in clinical medicine. *Gastroenterology* 46,788-791 (1964).

Foreword

The thought that led to this book began in John Duggan's work for the Repatriation Medical Authority in Australia. That Authority produces Statements of Principle which form causation templates for war veterans' compensation claims in Australia. These templates provide a definition and list of contributory causes for all diseases based on the current epidemiological literature. They are legally binding on all decision makers in veteran's compensation and are made at two standards of proof (i.e. Reasonable Hypothesis and Balance of Probabilities), which apply on the one hand to active war service and the other to non-warlike service. The need to meet the rigours of producing a legal document and the social purpose of providing generous compensation terms to war veterans produced an environment in which the five academic members of the Authority had to view the epidemiology in a fresh light. The multiple interactions of many factors responsible for all diseases became highly relevant to compensation claims. The need to make judgements at two different standards of proof on the same body of evidence focused attention on both the quality of evidence and the whole body of evidence and the process of scientific decision making itself. John Duggan's experience, wisdom and good humour were much appreciated in this process and will infiltrate this book to the readers' pleasure.

Professor Ken Donald
Head, School of Medicine
University of Queensland
Brisbane, Australia

Acknowledgments

This book would not have seen the light of day without the stimulus and assistance of many talented people. It had its genesis in the meetings of the Repatriation Medical Authority, a body set up by the Australian Parliament to determine the causative factors for disease based upon "medico-scientific principles" for the determination of pensions for Australian ex-service personnel for disability attributable to armed forces service. Over the eight years one of us (JMD) served on the Authority it became evident that determining the causation of a condition requires detailed and exhaustive literature review followed by critical expert analysis of the data. It was clear that the standard textbooks were often not only inadequate but sometimes lacked scientific credibility. For most organ systems, no such single source of information existed. So arose the concept of a book on the subject. To the Chairman, Professor Ken Donald and the members, Professor Dick Heller, now Professor of Public Health in the School of Epidemiology and Health Sciences in the University of Manchester England, Professor Beverley Raphael, Professor John Kearsley and Professor John Kaldor and the very talented secretariat staff, especially its leader, Dr Alex Bordujenko for their knowledge of epidemiology and their diligent support we are grateful. To Professor Heller, with his profound knowledge of epidemiology and incisive approach to problems much is due.

Mrs Patricia Aguado, Chief Librarian of the Gardiner Library at the John Hunter Hospital and her staff opened many doors to ferret out obscure references, forever cheerfully. Mrs Mary Howard also of the library staff provided the indexing.

Our secretaries, Mrs Amber Charman and Mrs Allison Colyvas typed numerous drafts skilfully and with equanimity. We are both aware of the

debt we owe to those of our own teachers who taught us to question dogma and tried to inculcate a spirit of enquiry in us. We are reminded of a possibly apocryphal story of graffiti outside a University Philosophy Department – "Question everything" under which someone had scrawled "Why?".

JM Duggan
AE Duggan

Chapter 1

EPIDEMIOLOGY AND CLINICAL PRACTICE

Epidemiology may be defined as "the study of the distribution of a disease or a physiological condition in human populations and of the factors that influence this distribution" (1). The term, however defined, refers in principle to determining factors which influence disease processes in humans. Although often associated with causes it may also concern itself with factors which diminish disease such as dietary folate and colorectal cancer, fluoride and dental caries, coffee and diabetes, exercise and ischaemic heart disease. The epidemiology of diseases is not a topic that makes most clinician's eyes light up. Equally, progressively medical schools are emphasising the social milieu in which illnesses present (2). In many respects, the sociological aspects of illness represent a low power view of the environmental influences so potent in health and its disorders – diet, water quality, social class, employment, housing, income, education, occupational hazards. We believe that epidemiology has a role and can be interesting to students and practitioners at the bedside just as it is relevant to more senior clinicians. One of us is old enough to recall when bronchogenic carcinoma was recognised as developing epidemic status when no-one seemed to have noticed the nicotine stains on the patient's fingers.

The definition given has two words worth emphasis. Firstly, we are concerned with "humans" which distinguishes it from animal research, although there must surely be a corresponding discipline for animal workers of which the complex story of the sheep industry in Australia provides an example. The second word to emphasise is "population". Epidemiology is not concerned with the individual as in standard clinical practice but with the health of the group. This takes us to two aspects of epidemiological work:-

1. Examination of the distribution of disease frequency in the population. This can produce hypotheses about the findings of the work. This is then a descriptive study.

2. Analytical studies which test the hypotheses generated by the descriptive studies.

1. DESCRIPTIVE STUDIES

These use population based statistics on disease incidence, morbidity, mortality, hospital and health data; anything dealing with disease or disability in the population.

2. ANALYTIC EPIDEMIOLOGY

This utilises data on patient characteristics, residence, occupation, lifestyle, leisure activity, diet, medication, the whole gamut sometimes of human existence. Such data are used to seek relationships between health statistics and population factors testing hypotheses of linkage and causation.

A characteristic so often seen in applying epidemiological data to clinical situations is the entrenched conservatism of practitioners and populations in respect of the utilisation of the findings. There are two outstanding examples which have quite significant historical importance, apart from their medical significance. In 1535 Jacques Cartier rediscovered Newfoundland and, as was usual, his crew was devastated by scurvy. Hakluyt in his "Principal Navigators" in 1600 described how a decoction of a local tree, probably the spruce fir, as advised by the local Iroquois Indians, provided a rapid cure. Yet the conventional wisdom could not accept a nutritional cause for the disease. James Lind in 1747 took 12 identical cases of the scurvy and gave each pair a purported remedy. Only the pair given two oranges and one lemon a day dramatically recovered. Published in 1753, it was ignored until James Cook utilised this knowledge for the triple circumnavigation of the globe and provided preserved cabbage throughout the voyage for his crew. He was the first navigator to escape scurvy and to discover eastern Australia and much else of the then unknown world and altered the course of history. Yet, only after 42 years could the Lords of the Admiralty be induced to prescribe a source of vitamin C for British sailors. Doubling the effective strength of the Royal Navy, this played a major role in Britain's defeat of the French navy in the Napoleonic wars.

Another example is of cholera which, present in Bengal for centuries, migrated to Europe and reached England early in the 19th Century causing 5500 deaths. It struck Sunderland in 1831 where a young medical apprentice, John Snow, observed the disease then attacking the coal miners particularly. He rejected the current paradigm of a 'miasma' or bad air as

the cause and concluded that diarrhoea with faecal contamination, particularly in coal mines, unwashed hands and shared food was causative. By 1848 he was a well regarded general practitioner in London when an explosive outbreak occurred. Snow, recalling his earlier experience still thought that the disease was spread by faecal contamination of water. He found that 73 of the 83 deaths in three days were in people living near and using the Broad Street pump. Eight of the other ten occasionally used the pump. After a prolonged debate he convinced the Board of Guardians of the local parish to remove the handle from the Broad Street pump. They did. The epidemic ceased. Thirteen years later, although he had achieved sufficient status to anaesthetise Queen Victoria for the birth of her last child, his discoveries had failed to achieve recognition leading to a tragic irony. The Queen's residence, Windsor Castle, still drew its water from the Thames into which it also emptied its sewage, leading to Albert the Prince Consort acquiring typhoid fever from which he died.

Both these examples of what could be called Public Health in action can be regarded as Applied Epidemiology. However, both illustrate two poorly appreciated aspects of the process. Firstly, those who like Lind and Snow, produce a new paradigm may have to face the apathy of the conservative medical and civil authorities (3). Secondly, epidemiological observations when put appropriately into practice may have major and lasting repercussions.

We may consider some aspects of the quality and type of evidence that epidemiological studies produce. *The US preventive services taskforce : Guidelines for Evidence* are valuable and provide a hierarchy to evaluate the quality of study design, 1 being the best and 5 the least potent.

3. ANALYTIC STUDIES

1. Intervention Studies
   - 1a. Randomised Controlled Trial
   - 1b. Controlled Trial
2. Observational Studies
   - 2a. Cohort-Prospective
   - 2b. Cohort-Retrospective
3. Case Control Descriptive Studies
4. Population (Correlation)
5. Individual
   - 5a. Cross Sectional Surveys
   - 5b. Case Series
   - 5c. Case Reports

We are influenced by these but a major influence in our evaluation is the Bradford Hill criteria (4).

1. Strength of Association
2. Consistency
3. Specificity
4. Temporality
5. Biological Gradient
6. Plausibility
7. Coherence
8. Experimental Evidence
9. Analogy

Such criteria are an essential tool in evaluating epidemiological evidence in a field so fraught with associations, bias effects masquerading as causes. To those who believe that epidemiology is an esoteric subject, largely residing in the computers of academics who have never seen a real patient for years, we extend our sympathies and particularly to their patients. Too often we have seen patients with Crohn's disease whose physicians have not realised that smoking is impairing the patient's response to their ready administration of potent but dangerous drugs.

One of the major influences in the rise of epidemiology and particularly analytical epidemiology has been the fantastic growth of computing capacity in the last several decades. In one of the key books of an earlier era, J N Morris' *Uses of Epidemiology* (1964) neither logistic regression nor meta analysis appear.

We have interpreted our role as one to bring to clinicians and to those concerned with health promotion an account of the current knowledge of the factors concerned with the development of disease, with some emphasis on the critical interpretation of the relevant literature. We have also tried to highlight some of the gaps in knowledge; if this stimulates young minds to attempt to fill those gaps, our joy will be unbounded. In a sense we are promoting Evidence Based Medicine, not in patient management, but in population management, with better management of the individual patient an important by-product. For example, if smoking is accepted as a significant promoter of peptic ulcer and Crohn's disease, then the clinician has a responsibility to promote smoking cessation as part of a patient's management plan, meanwhile supporting Public Health endeavours to reduce smoking in the community.

References

1 Lilienfield A M. *Foundations of Epidemiology*. (OUP, New York, 1976).
2 Susser M W, Watson W. *Sociology in Medicine*. 2nd Edn. (OUP, London, 1971).
3 Kuhn T S. *The Structure of Scientific Revolutions*. 2nd Edn (University of Chicago Press, Chicago, 1970).

4 Bradford Hill A. The Environment and Disease: Association or Causation. *Proc. Royal Soc. Med.* **58**, 295-300 (1965).

Chapter 2

SOME EPIDEMIOLOGICAL CONCEPTS

As we discussed in the previous chapter, epidemiology is the study of the distribution and determinants of disease frequency in human populations. To do this effectively we need some mechanisms to count events of interest and to observe populations.

These methods need to be applied to the two main approaches to epidemiology:

1. Descriptive studies: Studies that examine the distribution of disease frequency in populations to develop a hypothesis as to the cause of the disease, and
2. Analytical studies: Studies that test these hypotheses by reviewing personal characteristics of patients observed or, exposure to items of interest among individuals within the groups observed.

Descriptive studies use population-based statistics of disease occurrence, survival and mortality. Disease registers often provide important information. Two important concepts are:

Incidence - the number of new cases during a given period of time divided by total population at risk and

Prevalence - the number of existing cases of a disease divided by the total population.

Measures of exposure are often broad and subject to confounding or interfering factors. Before we conclude any associations between a disease and its possible determinant it is important to consider any factors that may influence study results.

Analytic studies are either observational or interventional. Clinical medicine relies heavily on interventional studies often in a randomized format to assess the value of proposed new interventions.

For the epidemiologist analytic studies are the bread and butter of epidemiology with two valuable study designs: case control studies and

cohort studies. Case control studies select subjects on the basis of whether they do (cases) or don not (controls) have a particular disease. The groups are then compared for the proportion having a history of exposure or characteristic of interest. A classic example is whether lung cancer patients have a more significant history of smoking than non lung cancer patients. Case control studies have a number of advantages. They are particularly suited to studying rare diseases because they allow the investigator to collect together an adequate sample size. They allow multiple exposures to be assessed and in the early stages of our knowledge of a disease may come up with some interesting surprise associations!

Subjects in cohort studies are classified according to their exposure to a particular factor (exposed or unexposed) and then followed for a specific period of time to determine the development of disease in the exposed group. As is easily imagined this can be years. The Nurses Health Study continues to be an important source of data having enrolled 120,000 female nurses in 1976. Second yearly these nurses complete an extensive questionnaire about a number of demographic, lifestyle aspects of interest and their medical history. This has allowed the investigators to identify important associations between lifestyle factors and diseases such as high blood pressure and a variety of cancers.

1. EVALUATING CAUSE AND EFFECT RELATIONSHIPS

Making judgements about causality from epidemiological data involves a change of logic that addresses two major areas:

a) Whether the observed association between an exposure and the disease is valid.

b) Whether the evidence provided from all the sources available support a judgement of causality.

Let us consider each of these issues in turn.

2. IS THE ASSOCIATION BETWEEN EXPOSURE AND DISEASE VALID?

There are three common explanations for why this might not be the case:

- Chance
- Bias
- Confounding

2.1 Chance

One can imagine that the smaller the sample size the greater the variability in estimates. Not surprisingly, one way to reduce the risk of chance is a large sample size. This reduces variability and provides more reliability of any inference made about the data. One important method for coping with variability in studies is to use confidence intervals (CI). Confidence intervals take into account the sample size and the expected degree of variability and provide a range within which the total magnitude of effect is likely to fall. A 95% confidence interval is commonly used and indicates with a 95% confidence that the true value of the variable (e.g. mean) lies within the interval. In contrast the p value, a composite measure of the magnitude of the difference between the groups and the sample size is susceptible to these factors and so a true difference may not be detected with a small sample size and visa versa. By convention if the P value is used then if it is less than or equal to 0.05, meaning that there is no more than a 5% chance of an observed result being due to chance, the association between exposure and disease is considered significant.

2.2 Bias

Bias is anything or process that deviates the results from the truth. This can occur by the way we select patients, observe them or question them and influence the way they respond.

2.3 Confounding

Confounding comes from the Latin to mix together and refers to failure to separate the effects of two processes. A confounder must be associated with the exposure of interest and independent of the exposure be a risk factor for the disease. An example of confounding is to conclude that a vegetarian diet reduces risk of cancer when it may also be that vegetarians also do a number of activities that also reduce cancer risk such as exercise.

3. GENERALISABILITY

Generalisability is another important question to consider when reading the results of epidemiological studies, that is: Do these results apply to the population in general or the population I am dealing with? Obvious

examples where there may be no generalisability would apply in results of studies of men to women or of one culture to another.

4. JUDGEMENT OF A CAUSE – EFFECT RELATIONSHIP

Simply because there is a valid statistical association between factors does not imply any causality between the two factors. It may simply be due to chance! As discussed in the previous chapter, in epidemiology there are a number of criteria that can aid in the judgement concerning causality. It may be worth considering these factors more closely:

- Strength of the association. It should not be a surprise that the stronger the strength of an association the less likely it is due to chance, bias or confounding.
- Biologic credibility of the hypothesis. It is important to remember that a statistical association that does not appear credible at one point in time may eventually appear to be so once we understand mechanisms.
- Consistency of the findings. Put simply the more studies supporting a hypothesis the more like the hypothesis is correct.
- Temporal sequence. It is logical that the exposure of interest should precede the disease of interest but this may not always be easy to establish.
- Dose – response – relationship. The observation of a gradient of risk with the degree of exposure is always reassuring.

Quantifying any association between exposure and disease is an important epidemiological tool. Relative risk calculates the size of the association between exposure and disease and the likelihood of developing the disease in the exposed group. It is the ratio of the incidence of disease in the exposed group divided by the corresponding incidence of disease in the unexposed. A relative risk of 1 indicates that the incidence of disease is the same in both the exposed and unexposed group. In case control studies cases already have the disease and so one can only calculate the odds of exposure among cases and controls. For example people who smoke may be shown to have a higher risk of pancreatic cancer or some of the other gastrointestinal diseases we are about to encounter. When we use the odds ratio we often calculate the corresponding 95% confidence interval to give with 95% confidence the upper and lower limits of the risk. Thus if we could come up with the estimate of relative risk of myocardial infarction with use of the oral contraceptive pill as RR (95%CI)= 0.7 (0.5-1.2). As the estimate includes 1 that is an equal risk in the exposed and unexposed groups the association is unlikely to be significant.

A word needs to be said about Attributable Risk, a measure known to epidemiologists but only now coming into the clinical literature. A simple example will illustrated. Suppose that in a population the prevalence of smoking is low and the relative risk of smoking for lung cancer is low. Then, in that population, only a small proportion of cases of the disease can be attributed to smoking. Conversely, if the prevalence of smoking is high and the relative risk is high, then a much larger proportion of the disorder can be attributed to smoking.

Put mathematically

If r = relative risk of lung cancer in smokers compared to non smokers and,

b = background prevalence of smoking in that population,

then Attributable Risk (AR) = b (r-1)/b(r-1)+1.

AR can be expressed as a percentage by multiplying by 100. It may be made to deal with multiple factors; standard errors and confident intervals can also be calculated. It also needs to be noted that as originally defined it is the "maximum proportion of lung cancers attributable to cigarette smoking". (See Lilienfeld, *Foundations of Epidemiology* NY, OUP.,1979).

Logistic regression is a very complex computational method of separating out the significance and strength to be attached to factors which may be contributing to an end result. If we consider various elements of diet, age, socio economic factors, duration in a condition like bowel cancer it can attack significance and strength to these factors.

Background Reading

Last J. *A Dictionary of Epidemiology.* 4th edn (OUP, Oxford, 2001).

Hennekens CH, Buring JE, (eds.) *Epidemiology in Medicine.* (Little, Brown and Co., Boston, 1987).

Chapter 3

GENETICS AND THE GUT

At first glance there appears to be little role for the epidemiological aspects of hereditary disorders, especially given that such disorders represent a minor part of the load of GI diseases in most communities. However, it will become evident, with modern techniques of genetic and molecular biology, that genetic factors play a significant role in the natural history of an increasing number of disorders. However, it must be recognised that, ultimately, all cancers have a genetic basis, with cancers representing the end product of complex interaction between environmental influences and specific genes of the subject. It is proposed to discuss a number of disorders that are important, both from their frequency, and mortality and our developing knowledge of their genetic background. For those seeking more detailed information, large genetic texts are necessary (1).

1. INFLAMMATORY BOWEL DISEASE

This consists of two similar but distinct entities, ulcerative colitis (UC) and Crohn's disease (CD), both frequent, important in terms of morbidity and mortality and for which genetic and environmental factors are intertwined in causation. The clinical factors are considered elsewhere; here we will emphasise the genetic aspects.

1.1 Racial and Ethnic Influences

For several decades it has been recognised that these are major influences; IBD was more frequent in Jews of Ashkenazi origin than in their fellow Americans, there was a marked north / south gradient in Europe and Caucasians had a higher incidence than non-Caucasians of both CD and UC

and in the U.S. there were higher rates in whites than blacks and Hispanics with Asians lower still. It is also evident that there is a strong familial tendency in both diseases with a 10-30 fold increase in the rate in siblings of patients. UC is increased in relatives of UC patients and CD in relatives of CD patients but CD and UC may co-exist in families more often than dictated by chance. It is also evident that familial aspects are stronger in CD (25%) than in UC (17%). In southern Israel IBD is nine times more prevalent in Jews than in non Jews. Whilst these differences may be attributed to environmental influences, several observations indicate a genetic basis; an increased prevalence in twins, absence of spouse concordance and concurrent disease in relatives widely separated geographically. Nonetheless, a straight Mendelian inheritance is excluded on the basis of these observations. Equally, detailed analysis excludes the proposal of a polygenic model where those with 10-15 specific genes develop CD, those with fewer develop UC and those with relatively only few of these genes develop UC rather than CD. The latest data indicate that the genetic basis for IBD is in the interaction of two or more major genes with genetic heterogeneity for UC and CD. This leads to the conclusion that IBD is a heterogeneous group of diseases with each subform an oligogenic disease due primarily to the interaction to a limited number of genes with or without a minor contribution from modifying genes (2).

As indicated elsewhere there are a number of epidemiological factors that influence the development of IBD such as smoking, diet, the oral contraceptive pill and appendicectomy. So these disorders have both genetic and environmental causes. A novel explanation for the relatively high frequency of the susceptibility alleles with their major roles in morbidity and mortality has been presented (3). A simple explanation is that they are due to mutations. However, their high morbidity and mortality would necessitate a very high mutation rate to compensate for the loss of life and reproductive capacity from them, one much higher than the estimated $1/10^5$ to $1/10^6$ rate per gene locus in humans.

Another explanation, the founder effect is also excluded. This concept is that a mutation occurs in an individual and is handed down to a high percentage of the descendants. While applicable to isolated communities and certain ethnic groups, it is invalid for IBD with its high frequency and wide distribution. The remaining explanation is best, namely that carriage of the IBD gene also carries a selective advantage. Examples of this in gastroenterology already exist; the protection against cardiovascular disease and hypertension in DU and the protection against iron deficiency given by haemochromatosis. The concept is that the possession of a disease producing gene and a benefit ultimately leads to an equilibrium between benefit and ill effect. In a case of IBD, the presence of the IBD genes