EDITORIAL

The epidemiology of life stress

Different contemporary societies, different historical periods in the same societies, and different strata within the same societies, produce different patterns of health disorders. At the ecological level, there can be no dispute that particular social structures generate particular patterns of health disorder. To put it another way, the distribution of health in a society is one dimension of the social structure and an indicator of its nature and impact on people's lives (Susser & Watson, 1971).

When we turn to the individual level, however, and to the question of how the impact of the social structure is translated into pathology, much that is said must rest on assertions of authority and faith. It is about 40 or 50 years since the 'stress' hypothesis gathered momentum and began to displace the fashionable hypothesis of that time, which attributed a gallimaufry of obscure ills to foci of infection. Many excellent sets of teeth and many pairs of tonsils were sacrificed to that hypothesis. We can count ourselves fortunate that the stress hypothesis does not require surgical intervention.

The starting point for social scientists is concern with society, and hence with total populations, both the healthy and unhealthy, rather than with disordered segments of populations. They thereby gain breadth. Social scientists extend our vision also because they tend to begin with the independent variable—a social factor—and seek out its many effects, rather than begin with the dependent variable—a health disorder—and seek out what impinges on it. To start with health disorders is a narrowing view that is the norm in medicine, since its concerns arise there. Multifaceted society leads social scientists quickly and naturally to multivariate paradigms, rather than to the specificity of one-to-one relationships.

Much of my own endeavour has been to bring the concepts and the methods of the social sciences into medicine and epidemiology. I shall here reverse myself and try and bring out for social scientists the notions and emphases of epidemiologists. One of these is the notion of the triad, agent–host–environment, and the need, in analysing the problems of health disorders and in constructing research designs, to segregate these three interdigitating elements. This triad is entrenched in epidemiology. Indeed, the triad appeared at last to have become an impediment to flexible multivariate analysis, to the extent that the subject needed to be freed from its tyranny.

In counterpoint, the literature of 'stress' disorders shows that many investigators could have benefited from familiarity with the concept. There has been a lack of differentiation among the components of the triad: some studies concentrate on agents, namely stressful events; others on the environment, namely situations; while the characteristics of the host have been virtually ignored, and we know little of susceptibilities and resistance.

In classifying agents, researchers conscious of the triad might less often have lumped stimuli indiscriminately together, irrespective of force and direction. Stress disorders can be construed in the manner of Hans Selye as the intermediate result of some final common pathway, analogous to inflammation, which is a general pathological response to many agents and the antecedent of a variety of specific manifestations of disorder. Understanding of inflammation was much advanced when the agents were sorted one from another: chemicals, heat, ultraviolet and ionizing radiation, bacteria, viruses, protozoa and multicellular parasites all cause inflammation. In the field of stress, it is clear that specified events yield more convincing evidence than do undifferentiated ones: 'loss' emerges as the particular antecedent of depression (Paykel et al. 1969).

In describing hosts, researchers conscious of the triad might have paid more attention to the varying susceptibility and immunity of individuals with different constitutions and experiences, including variability denoted by 'simple' indices like age, sex and other statuses and roles. No disease is uniform in all these respects.

1 Address for correspondence: Dr Mervyn Susser, Gertrude H. Sergievsky Center, Columbia University, 630 West 168 Street, New York, NY 10032, USA.

0033-2917/81/2828-5620 $01.00 © 1981 Cambridge University Press
In specifying environments, it might quickly have become obvious that the action of any agent or stimulus cannot meaningfully be considered free of context and situation. A social variable like population density will have different meanings for crowding and its effects when measured in terms of persons per room, or per household, or per area. Within families the likelihood of common respiratory infections, for instance, is quite altered by the presence of children of school-going age. Likewise, with recurrent psychotic episodes, family environment must be expected to cushion or exacerbate the impact of stressful events (Brown et al. 1972).

Secondly, the problems of disease that confront epidemiologists present themselves in many guises, in a great variety of circumstances, and with varying degrees of urgency. Health disorders may be acute, recurrent, or chronic; they may be mild or severe; they may be of abrupt or insidious onset; they may be distinctive or shade gradually in a continuum in which it is difficult to tell health from disease; they may be rare or common. To cope with this variablity has required many approaches to research design. Epidemiology is best pursued with the largest possible arsenal of designs and strategies at hand. Social epidemiology is not less in need of flexible and complementary approaches.

Thirdly, as noted, the concerns of medicine and of epidemiology arise from the primary source of manifest health disorders and pathology in individuals. If the response to these problems is to be effective, then the realities of the biological, genetic and demographic substrate which carries the imprint of disease cannot be bypassed or ignored.

Fourthly, the history of the explosive advance of biomedical knowledge in the nineteenth and the first half of the twentieth century was the history of a search for highly specific causal relationships, to the point of gross distortion when single causes were sought for each disease. Nonetheless, this history has its dialectic: it is also the history of the continuous refinement and specification of relationships, of the dogged elimination of alternative hypotheses, of the increasing specificity and ‘testability’ of hypotheses, and the resultant strengthening of inference.

The best and most active researchers are not in need of sermons from the side-lines: they are already testing many of these ideas. The care that has gone into recent reports and the rigorous critical effort that characterizes recent reviews augurs well (Dohrenwend & Dohrenwend, 1974; Rabkin & Struening, 1976; Brown & Harris, 1978a, b, 1979; Tennant & Bebbington, 1978; Shapiro, 1979). But a reading of the mass of the literature suggests that discussion is needed nonetheless. The themes fall under the heads of (A) Instruments, (B) Design, (C) Analysis and hypothesis testing, (D) Control and prevention of pathology.

INSTRUMENTS

Reviews by others of efforts to measure the cumulative impact of stressful life events point to serious shortcomings of a popular instrument (Dohrenwend & Dohrenwend, 1974; Rabkin & Struening, 1976). The readily used approach of Holmes & Rahe (1967) produced a great spurt of studies about the impact of life events. Scientific advance has regularly followed from the development and accessibility of new techniques, whether the microscope or the randomized clinical trial, and now in the study of stressors too we may be hopeful. Typically with new techniques a hundred flowers bloom, but much weeding and culling must usually be done to refine them before notable advance can follow.

Researchers have come to recognize and correct for a number of obvious problems with the technique. If we aim to separate cause from consequence, a single scale that combines events, many of which are not certainly antecedent to the manifestations under study, is disqualified. For instance, many of the supposed stressors are transitional events. For any one such event, say divorce or any other, the task of separating a stress reaction to an event from self and social selection for exposure to that event has proved difficult. The history of studies in the area makes this plain. Divorcees may not only become divorced and confound inference because they are physically or mentally sick; like the single, they may remain unmarried for that reason.

A further intractable problem in establishing the criterion of time-order to secure causal inference is the possibility of predisposition in the host. Predisposition conceived as an independent variable and controlled from the outset is manageable; predisposition conceived post hoc as a latent manifesta-
The instruments in use have obviously needed considerable refinement. Magnitude, direction and familiarity of stressors should not be neglected. Categorization of events can help to define both host susceptibility and context. Susceptibility varies with stage in the life-cycle which differentiates certain universal transitions, *rites de passage*, and normal role changes as incremental or decremental. Context and situation are bound to be important. Animal experimenters teach us that the social context determines the impact of a biological insult; for instance, nutritional deprivation has no detectable effects on associative learning in rhesus monkeys or in rats when they are not socially isolated (Zimmerman *et al.* 1975; Levitsky & Barnes, 1970). Finally, since psychosocial stressors of necessity are mediated by the psyche, we need to find the means to take into account the perception or denial of stress by the subject, as Brown (1974), Cobb (1974) and others have emphasized.

These are problems of logical construction. Dangers also arise from the fundamental technical difficulty of establishing validity. The basic problem which these instruments have shared with many other social instruments is their weakness as measures of reality. The simple danger is that if the instrument fails to measure the hypothetical causal variable, no association (causal or otherwise) can be demonstrated. One must, therefore, be ready to accept the social costs of a false negative result.

Researchers have laboured with determination and good effect to perfect population survey instruments for the valid and reliable measurement of mental disorder, the dependent variable in a majority of previous studies (Srole *et al.* 1962; Leighton *et al.* 1963; Weissman *et al.* 1977). Equal attention needs to be given to the independent variable. Brown and his team have put effort and care into developing an instrument that takes account of both circumstances and perceptions in assessing events listed in a standardized dictionary; the gains in sensitivity must compensate for losses in simplicity and portability. Reliability studies are under way (Neugebauer, 1980). Compared with reliability, however, the validity of measures of stressors has had little attention. An effort has been made in one study to test whether the events reported did in fact happen, and as many as 80% could be confirmed. This indicates a degree of validity in measuring exposure to events, but whether they are in fact stressful events is another question. The discovery of an association tends to be taken as in itself a validation of an instrument. Where events that carry much freight besides affective stress are under study, this is not a safe assumption. For example, a change in marital state often implies marked changes in income, in diet, in exposure to infectious disease, in life style and habits, as well as in emotional state.

Validity needs to be tested, as difficult a task as this may be. For this purpose one might look to a marriage of social and psychological stress indicators with physiological indicators. In the almost equally difficult measurement area of nutritional intake, for example, measures are beginning to achieve a degree of validity in terms of changes in body weight and the excretion of nutrients in the urine. Stressors are surely mediated through the central nervous and the endocrine systems. Thus, several physiological indicators responsive to short-term emotional change are available for testing as potential indicators of longer-term stress. A start has been made by Cobb and his colleagues (1973), Theorell (1974), and others (Stevens, 1959).

With regard to measurement of the dependent variable, the problems are more tractable. Mental disorder is the most obvious and likely, if not the sole, outcome of stressful experience. In psychiatric epidemiology, where the definition of a case has been a central problem for decades, a good distance has been travelled. The unidimensional scales used in many prevalence surveys created problems that resemble those of undifferentiated lists of stressful life events. The differentiation in the diagnosis of mental state attained through the research of the past decade has made a distinct advance on the unidimensional approach (Wing *et al.* 1974; Spitzer *et al.* 1975). Other workers have begun to exploit and develop these diagnostic instruments for population survey work. So long as we have had to acknowledge that a case of psychiatric illness was defined with difficulty, we have also had to acknowledge that we could not be precise about the effects of exposure to life stresses. We should do better on this score in the future.
DESIGN

Fortunately, social and psychiatric epidemiologists are no longer wedded to the prevalence survey as the only tactic for reaching out to populations. But a too ready commitment to cohort studies as the universal replacement for cross-sectional surveys, and the dismissal of case–control designs, would be an over-reaction. Major problems in the case–control design are to establish time-order, to recognize selective bias where there is a long time interval from exposure to manifestation, to select truly comparable controls, and to recognize the multiple outcomes of a single type of exposure.

In case–control studies of the effects of life events, the outstanding unsolved problem is the bias that is likely to inhere in the post hoc reports of events by the cases: their recollection is coloured by the outcomes that have befallen them and made them cases. A notorious, unwitting example was the demonstration of an association of Down’s syndrome births with ‘stressful’ pregnancies. Another source of bias, contamination by the investigator’s subjective judgement when either the potential stressors or the hypothesized effects are assessed or scored, is amenable to control by standard techniques which keep the assessors of the effects blind to knowledge of stressors and vice versa. In reports of even the best studies it is not always clear that such precautions have been rigorously applied.

On the other hand, the case–control design is efficient and parsimonious; it allows the estimate of the odds ratio as an equivalent of relative risk, even though it cannot yield a direct estimate of attributable risk; it permits the simultaneous study of several independent variables; and often it offers the only possible way to study rare diseases (MacMahon & Pugh, 1970; Susser, 1973; Fleiss, 1973; Ibrahim, 1979).

The cohort design is superior in theory with regard to establishing both time-order and the starting population. There is no problem of retrospective bias in eliciting exposure, and observer bias in making observations of outcome is amenable to control. One can also recognize the multiple outcomes of a single type of exposure, such as a psychosocial stressor, and the design is likely to be superior for the study of rare causal variables, for example war-time bombing or famines (Neel & Schull, 1956; Stein et al. 1975). In practice, the design may be less satisfactory than in theory. During years of laborious prospective observation, the nature of the problem may undergo change with the advent of new knowledge; expense may restrict observation to a single cohort or generation without the advantages of quick replication; rare outcomes may be impossible to study. Cohort studies tend to be most advantageous when cohorts can be reconstructed from existing historical data (Stein et al. 1975), or when they can exploit existing longitudinal data sets (Vaillant, 1979).

A cohort design, moreover, is not always essential to establishing time-order. Without benefit of cohort designs, Goldberg & Morrison (1963) fixed the phenomenon of downward drift in schizophrenia by comparing the social class of young schizophrenic men with that of their fathers recorded on their birth certificates. The difficult question of the effects of the loss of a spouse has also been clarified by a succession of studies using other types of research design. In trying to demonstrate assortative mating, Karl Pearson and his colleagues at the turn of the century found evidence, from studies of parish registers and tombstones, of an association of age at death among spouses (Biometrika, 1903). In the early 1940s, Anthony Ciocco (1940) followed this lead, meaning to counter Pearson’s hereditarian bias. He linked the death certificates of spouses in a Maryland county, and demonstrated associations among them, not only for age at death but also specific for several causes of death. Later, Kraus & Lilienfeld (1959) showed that in national statistics the widowed, and especially the young widowed, had excessive death rates from several causes.

These early studies laid the ground for the testing of specific hypotheses. Thus, MacMahon & Pugh (1965) demonstrated the role of widowhood in suicide in a case–control study. They examined the death certificates of suicides and controls for the status of their spouses and, when deceased, the date of the loss, and showed a clustering of suicides following recent bereavement. Parkes (1964) indicated the possible further effect of recent widowhood on entry to psychiatric care by comparing the bereavement experience of Maudsley Hospital in-patients with that expected from the mortality
Editorial: Epidemiology of life stress

rates of England and Wales applied to the age and sex distribution of the spouses of patients. This study was vulnerable to confounding, because lack of social support is associated with both widowhood and hospital admission, and because the national comparison was an imprecise control. Stein & Susser (1969) confirmed the result, however, by demonstrating the clustering of recent bereavement among inceptions of psychiatric care in a community register of all referred psychiatric disorder, as compared with all chronic psychiatric disability and with a local census population.

What is needed is a flexible approach and an arsenal of methods. Different approaches vary in their utility in different situations. They should be seen as complementary to each other. If we are fortunate, a variety of approaches will give a result the force of replication. If we are less fortunate and the results are contradictory, incoherence helps either to detect flaws in method or to generate new theory.

The Framingham study of coronary heart disease is one of the great exemplars of the cohort design (Dawber & Kannel, 1963) and it too has been put to use, although belatedly, to study the effects of psychological factors in coronary heart disease (Haynes et al. 1978a, b, 1980). But the case-series and time-series studies which preceded the Framingham study (for instance, the decline in heart disease in a war-beleaguered and hungry Europe in the early 1940s) yielded the hypotheses to be tested. Many subsequent case-control studies have helped us follow leads generated by the Framingham study, and now the stage is set for efforts to control pathology through intervention studies. This is a natural sequence in epidemiological investigation.

In another example taken from one of our own series of studies, we sought to understand the socializing influence of family, class and culture on development. For this purpose we chose the symptom of enuresis as likely to be sensitive to such socializing influences. In order to explore this large range of influences, we found it necessary to carry out many complementary studies: a community prevalence survey established the distribution, variation and natural remission of the condition and the true scale of the problem; intensive family interviews with a case-control design differentiated the effects of socialization on sphincter control within family types; studies of populations in special institutions (reform schools, children's homes, foster homes, day-care centres) in which cohort, historical cohort and case-control designs were used to demonstrate that disruptions of socialization had adverse effects on the frequency of enuresis, and that restoration to family settings had reparative effects (Stein & Susser, 1967). In this series an intervention study, in the form of the first controlled trial of conditioning treatment, came early and with a practical objective rather than at the end, but that study too strengthened the plausibility of the hypothesis that nocturnal sphincter control could be seen as a learned behaviour. For excellent examples of the complementarity of several designs in the solution of specific problems, one can point to the studies of Cobb, Kasl and their colleagues on rheumatoid arthritis (Kasl & Cobb, 1969) and of the late John Cassel and his colleagues on the effects of culture change (Cassel & Tyroler, 1961).

ANALYSIS AND HYPOTHESIS-TESTING

Rigor in setting up hypotheses, and close scrutiny of underlying assumptions, accelerates the acquisition of new knowledge. This testing approach has been too little practised. For example, the effects of social mobility and resulting inconsistencies of status have long been favourite topics in the social sciences. Yet almost the only consistency that exists among the results of studies of the stressful effects of status inconsistency and social mobility is their inconsistency. In part this muddle arises from weak analysis and inference and, especially, from failures to control for statistical interaction and for status levels at the starting point of change (Blalock, 1967). While these problems have come to be better understood, ecological and individual levels of study and analysis are still often not differentiated. For instance, it is inviting fallacy to amalgamate without qualification, as examples of results of culture change, the studies of Cassel & Tyroler on individual adaptation to migration with their studies of the changes that occurred in populations as society changed around them in a given area (Cassel, 1971, 1974).

With regard to hypothesis testing, we have enough knowledge to generate concepts and broad
theory, but too little knowledge at the specific level to enable us to examine the validity of such broad theories across many conditions. To achieve this we shall need to study specific stressors and specific outcomes. The Holmes & Rahe approach through a compendium of stressful events, it should be noted, is essentially a test of just one plausible hypothesis: namely, that stress is a response to the cumulation of events. There are as many other possible hypotheses as there are specifiable types of stressors and specific outcomes. The replication of studies addressing the same broad theory seems endless. A close look at what has emerged from these explorations, combined with disciplined review of biomedical knowledge for given disorders, must yield new designs and more answers.

Strong inference demands the effort to eliminate alternative explanations. For example, with regard to the inconsistencies in the social class distribution of coronary heart disease that have been apparent for the past three decades, the difference in the spread of occupations nationally and within organizations has been no better an explanation than the misclassification of disease at death (Lilienfeld, 1956; Lehman, 1966; Antonovsky, 1968). Little effort was made to test these hypotheses against each other. However, neither of these hypotheses is the most likely explanation. It is slowly becoming apparent that secular change in the disease is probably affecting different social strata in different ways (Marmot et al. 1978; Susser & Watson, 1971). Such competing explanations can be examined in study designs that are capable of eliminating them.

CONTROL AND PREVENTION OF PATHOLOGY

Most observations relevant to controlling the effects of stress would require individual behaviour change. Aside from the difficulty, and sometimes the questionable desirability, of inducing such change, much current life stress research can add little to our ability to control health disorders. The demonstration that the cumulation of a wide array of undifferentiated stressful events can generate stressful states does not point to the means of prevention or treatment.

On the other hand, the study of specific major events can point the way to prevention. If the immediate aftermath of widowhood puts a person at risk for suicide (MacMahon & Pugh, 1965), death from cirrhosis of the liver (McNeil & Kelsey, 1974; Parkes et al. 1969) (presumably caused by alcoholism) and other causes (Ciocco, 1940; Kraus & Lilienfeld, 1959), as well as for entry to psychiatric care (Parkes, 1964; Stein & Susser, 1969), then a target group has been defined and various interventions can be tested.

Similar potential for intervention resides in demonstrating the effects of lack of social support, and possibly even those of malfunctioning social networks. Germane theory has long been available, from anthropology and sociology (Bott, 1957; Townsend, 1957; Young & Wilmott, 1957; Cobb, 1976; Susser & Watson, 1971), and attempts to apply it have begun (Raphael, 1977). The availability and effectiveness of the family networks of children can have crucial effects on their subsequent careers and behaviour (Goldfarb, 1943; Bowlby, 1952; Gregory, 1965, 1966; Stein & Susser, 1960; Skeels, 1966; Rutter, 1966). Recent attempts to apply theories of social support suggest that supportive networks may protect the health of adults as well as children (Nuckolls et al. 1972; Berkman & Syme, 1979). For specified vulnerable groups, it is not fanciful to think that health and social agencies could intervene effectively.

The detection and control of harmful environments is a classical public health strategy. The strategy can be applied to the social as well as to the physical or the biological environment. It can be objected that a health agency, which is necessarily an arm of established society, cannot itself be an instrument that overturns or even modifies the social structure created by that society, its productive relations, and its political forms. This objection does not rule out planned change within that structure of an order sufficient to benefit health.

It may be well in this instance to change the paradigm from the individual to the ecological level. The most powerful forces that effect change in the disease patterns of populations clearly operate at this level. If we take history as our guide, we can be sure that to eliminate poverty, and indeed to restructure ways of life, will change the health of populations.

We need not be overweening in our ideas. The example of smoking, the major specific health
hazard demonstrated in the post-war era in developed countries, will serve. Successes in controlling smoking at the individual level seem minimal, and have rendered many health professionals pessimistic. At the ecological level, the picture is much more optimistic. There is little question that there has been a distinct decline in the smoking habit in successive cohorts, although the decline varies according to sex and social position (Hammond & Garfinkel, 1968; Moss, 1979). This is, I believe, the result of what amounts to a social movement that has grown powerful enough to counter the economic, social and psychological forces that have sustained the smoking habit. A social movement is constituted by groups that share social values and aim to transform the values of the society around them by example and private persuasion, and by education, political action and public policy: the groups themselves and the actions they pursue have both organized elements (with leaders, membership, and defined programmes) and unorganized ones (with informal interaction among partisans) (Herberle, 1967). In the light of this social movement model, epidemiologists might add a new set of social strategies for the control of pathology. Prevention and control constitute the ultimate business of health professionals. They must enlist a broad range of forces before they can hope to achieve them.

Mervyn Susser

Richard Neugebauer and Zena Stein were kind enough to read this script and enhance its rigor. Deborah Kayman helped with the references and Carol Giles with the typing. The paper had its origin in the Symposium on Psychosocial Epidemiology, organized by Saxon Graham for the American Sociological Association meeting in New York in 1976.

REFERENCES

Haynes, S., Feinleib, M. & Kannel, W. B. (1980). The relation-