What will a primary care led NHS mean for GP workload? The problem of the lack of an evidence base

Lone Lund Pedersen, Brenda Leese

Summary

Ongoing negotiations on the general practitioner contract raise the question of remunerating general practitioners for increased workload resulting from the shift from secondary to primary care. A review of the literature shows that there is little evidence on whether a shift of services from secondary to primary care is responsible for general practitioners' increased workload, and scope for making generalisations is limited. The implication is that general practitioners have little more than anecdotal evidence to support their claims of greatly increased workloads, and there is insufficient evidence to make informed decisions about remunerating general practitioners for the extra work resulting from the changes. Lack of evidence does not, however, mean that there is no problem with workload. It will be increasingly important to identify mechanisms for ensuring that resources follow workload.

BRENDA PRINCE/FORMAT

Changed discharge practices in the acute trusts means that general practitioners and other practice staff or community nurses are likely to be caring for elderly people discharged from long stay and acute inpatient hospital departments

National Primary Care Research and Development Centre, University of Manchester, Manchester M13 9PL

Lone Lund Pedersen, *research associate*

Brenda Leese, *research fellow*

Correspondence to: Ms L Pedersen. LPEDERS@fs1.cpcr.man.ac.uk

BMJ 1997;314:1337–41

Background

A primary care led NHS places emphasis on shifting the balance of care from the acute hospital sector to primary care.[1-3] This is only one of the many changes taking place in the NHS, and its effect on workload in primary care is potentially far reaching.

The trend for more services to be provided in primary care has been in progress for some time, principally as a result of technological changes and rising consumer demand. Furthermore, the 1990 general practitioner contract created financial incentives for general practitioners to replace some hospital based services with practice based provision, for example, by providing services such as minor surgery and chronic disease management for diabetes and asthma.[4] Fundholding created incentives for general practitioners to provide secondary care in their practices, for example, by using savings to invest in their premises and practice based facilities.[5] Further developments included in guidelines from the NHS Executive in 1993 allowed fundholders to use their budget to pay either themselves or other health professionals to provide a specified list of secondary care services.[6]

Other changes that had implications for general practitioner workload, such as long term policies to move care away from institutional and hospital settings towards care in the community, were re-emphasised in the 1990 NHS and Community Care Act.[7 8] In addition, the internal market encouraged trusts to change discharge practices by reducing length of stay in order to maintain or increase throughput, in response to financial pressures created by the purchaser-provider split.

The NHS Executive's publication which set out "a national framework for the provision of secondary care within general practice"[9] also had implications for general practitioners' workload. This trend has accelerated with the changes put forward in *Primary Care: The Future* and in *Choice and Opportunity*.[10 11] Although the reforms offer general practitioners a central role in the NHS, there are reports of increasing dissatisfaction,[12-14] coupled with substantial resistance to change, with general practitioners pointing to anecdotal evidence of increased workload as a barrier to further change.[15-21]

At the centre of the debate is whether general practitioners will carry out the work that arises from shifts in the balance of care without being given extra resources. In *Primary Care: the Future*, a precondition for taking a primary care led NHS further is to ensure that resources follow transfer of activities into general practice.[10] Remuneration is also at the centre of the negotiations about the general practitioner contract. The profession's most recent document sets out a

national definition of core services, which general practitioners are contractually required to do, as well as "non-core" tasks, which include many services with the potential to be transferred from secondary care and for which specific payment should be made.[22] This, in effect, is the profession's version of a new contract.

Searching for evidence

What evidence is there that a primary care led NHS will necessarily increase general practitioners' workload and create the need for compensation? We undertook a review to identify what secondary care services are now provided in primary care, and to summarise the evidence available on the impact on workload.

The published literature was searched, using several databases including Medline, Social Science Citation Index, and BidsEmbase. The search proved difficult: keywords relating to the interface between secondary and primary care are poorly developed. Keywords relating to shifts in the balance of care (see box) were identified from general reviews. The term "workload" did not catch many studies relating to the shift, since data on the impact on general practitioner workload was frequently a minor part of a wider study and was not dealt with explicitly. For these reasons, a central part of the search consisted of consultation with experts and searches of the reference lists of selected texts. Studies were included only if the design made it possible to distinguish between the workload generated from shifts in the balance of care and the workload that would have been in primary care in any case—that is, controlled comparisons and "before and after" studies. In all, we surveyed more than 200 studies; the list is given in the full report, which is available from the National Primary Care Research and Development Centre.

Schemes and developments aiming to shift the balance of care into primary care

Changed practices in primary care:
- Minor surgery performed by GPs
- Shared care schemes for chronic disease management
- Open access services
- Preoperative assessment and work up
- Follow up after surgery
- Minor injury units and accident and emergency

"In house services":
- Outreach clinics
- Directly employed specialists
- Direct service provision by GPs

Shifts to community based care and changed discharge practices:
- Reduction in long stay provision
- Reduction in acute provision; early discharge and day surgery
- Relocation of acute provision; hospital at home
- Relocation of long term patients to the community based care

Potential developments:
- GP led, continuing community based care
- Transfer from accident and emergency departments to general practice

Nature of the secondary-primary care shift

The 15 distinct areas in which a transfer of activities from secondary to primary care has been identified are listed in the box. "Shifted activities," for which only the location of care has changed, can be distinguished from "substituted activities," which also involve a shift in the type of healthcare professionals delivering the service. Workload implications for general practitioners arise only in the case of the substituted activities, but in many instances the additional work might be undertaken by other professionals within the primary care team.

Many general practitioners are now providing a range of services at their premises, rather than referring patients to hospital. Thus, most are likely to be directly involved in minor surgery and shared care for chronic disease management. Some general practitioners are also providing specialist diagnosis and treatment as an in house service. Those who have gained open access to, for example, surgical waiting lists, are providing preoperative assessment, and in a few places are providing postsurgical follow up of patients who are discharged to general practice rather than to outpatient clinics. In addition, the shifts to community based care and changed discharge practices in the acute trusts could have knock on effects for all general practitioners. Thus, general practitioners and other practice staff or community nurses are likely to be caring for patients discharged early from hospital, or after day surgery, as well as providing continuing care for severely mentally ill people, younger people with learning disabilities, and elderly people discharged from long stay and acute inpatient hospital departments. Developments such as general practitioners taking over routine work from accident and emergency departments and community based continuing care led by general practitioners are not, as yet, widespread. Consultant outreach and directly employed specialists are examples of shifted activities where the workload implications for general practitioners are minimal.

Evidence on implications for workload

Much of the literature on general practitioners' workload has arisen in the aftermath of the 1990 general practitioner contract and has focused on the impact on the contract's administrative requirements and increased participation in health promotion and prevention. Only 12 studies have provided evidence on the specific workload implications of secondary care services being shifted into primary care,[23-34] and only one study has addressed the issue directly.[23] This study examined the effect on workload of day surgery and the care required after discharge to the patient's home. The remaining 11 studies are all cost effectiveness studies that provide evidence which allows the evaluation of the effect on workload of four additional types of changes. Two randomised controlled trials evaluated the introduction of general practitioner led shared care for chronic disease,[24 25] and one randomised controlled trial evaluated postoperative follow up in general practice instead of in outpatient clinics.[26] Six studies[27-32] gave information on the potential knock on effect of schemes in which well organised community based services substituted for hospital inpatient provision for mentally ill patients. Of these, four randomised

Table 1 Overview of the effect of shifts in the balance of care on workload in primary care

Care type	Author and setting	Study design	Definition of workload	Effect on workload in general practice (extra workload)
Shared care based in general practice versus outpatient clinics:				
Diabetes	DICET (1994)[24] Aberdeen, Grampian	RCT	Number of consultations in general practice by patients under shared care during one year	Average consultation rate 2.6 per year†
Asthma	GRASIC (1994)[25] Aberdeen, Grampian	RCT	Difference in number of consultation in general practice between patients managed in outpatient and patients under shared care during one year	No difference
Postsurgical follow up in general practice versus outpatient clinics:				
General surgical patients	Florey *et al* (1994)[26] Dundee	RCT	Difference in number of consultations in general practice between patients discharged to GP follow up and patients followed up in outpatients in the first six months after discharge	Patients discharged to follow-up in general practice had 0.25* more consultations (on average 2.21).
Relocation of long stay provision: community care versus long stay hospital:				
General psychiatric patients	Hallam *et al* (1995)[31] North London	Matched prospective comparison	Percentage of patients discharged from long stay hospital who used general practice services in the first 12 months after discharge.	82%
Schizophrenic patients	Beecham *et al* (1995)[32] North London	Before and after	Percentage of patients discharged from long stay hospital who used general practice services 9-13 months after discharge	79%-89%
Elderly people with dementia	Knapp *et al* (1994)[33] North London	Before and after	Average cost of community residential provision per person per week accounted for by GP consultations.	On average £6.30 per patient per week
Relocation of acute provision: home based care versus standard hospital care:				
General psychiatric patients	Knapp *et al* (1994)[27] London	RCT	Difference in percentage of patients in home based care and patients having standard hospital care who used general practice services after four months and 12 months	Medium term: fewer of patients in home based care made use of general practice (36% *v* 54%); long term: no difference
General psychiatric patients	Burns *et al* (1993)[28] London	RCT	Difference in number of consultations in general practice between patients in home based care and patients having standard hospital care during one year	Patients in home based care had 0.7 more consultations (on average 8.3)
Schizophrenia patients	Burns *et al* (1991)[29] London	RCT	Difference in number of consultations in general practice between patient in home based care and patients having standard hospital care during one year	Patients in home-based care had 4.5 fewer consultation (on average 6.0)
Relocation of acute provision: early discharge of patients to day hospital versus standard inpatient care				
	Dick *et al* (1985)[30] Dundee	RCT	Difference in number of consultations in general practice between patients with early discharge to day hospital care and patients having standard care per month	Patients in day hospital care had 0.5* more consultations at three weeks after randomisation (on average 1.2), falling to 0.2 extra consultations at four months
Early discharge to hospital at home scheme versus standard inpatient care:				
Patients with fractured hip	O'Cathain (1994)[34] South Derbyshire	Prospective controlled comparison	Mean time spent with each patient during a period of 12 days by hospital at home team	Nursing time per patient:‡ Grade B nurse 12.67 hours, Grade G nurse 5.02 hours, Grade E nurse 1.74 hours
Reduction in acute provision:				
Increases in day surgery plus change of case mix from mainly minor procedures to intermediate procedures	Pill and Stott (1995)[23] South Glamorgan	Before and after	Percentage of patients having undergone day surgery consulting their GP during the first seven days after discharge. Exact time of consultation. Length of consultation	25% (22%) of patients had one or more contacts; 89% (90%) of consultations took place in normal surgery hours; less than 10 minutes of GPs' time was required for 82% (87%) of consultations§

RCT=randomised controlled trial.
* Significant at least at 5% level.
†Consultation rate in control group not given.
‡Input of general practice not given.
§The figures are given for the minor procedure sample. Figures for the intermediary sample are given in parenthesis. The difference indicate the effect of change in case-mix towards intermediate procedures.

controlled trials evaluated the introduction of comprehensive community based care, rather than standard hospital care, for acute psychiatric patients[27-30] and two evaluated schemes for comprehensive community based care to replace long stay hospital provision for psychiatric patients.[31 32] Finally, two studies indicated the resources needed for continuing community based care: one considered costs of general practice services for elderly patients with dementia discharged from long stay hospital to nursing homes,[33] and another assessing a hospital at home scheme.[34] Table 1 gives details of the information extracted from these studies.

The only scheme which clearly resulted in additional workload for general practitioners was discharge of patients with long term mental illness from psychiatric hospitals. Shared care schemes for chronic disease management for diabetes and asthma, and discharge of surgical patients to general practitioner follow up, had implications for workload, but these were minimal compared with the situation where no schemes exist, since patients visit their general practitioner regardless of whether they are attending outpatient departments. As predicted in the wake of the Audit Commission's publication of potential day surgery targets in 1992,[35] the knock on effect of increases in day surgery on general practitioners' workload was minimal.[36] Although the level of follow up required by day surgery patients varied between studies, this conclusion was reached by other studies as well.[37 38]

Figures from the workload studies have been used to model the impact of these changes on a typical week's work for general practitioners. Estimates were based on expected prevalence of the conditions in question in a typical patient list of 2000, and a mean consultation rate of three visits per patient per year.[39] Implementing shared care for management of

diabetes and asthma, long term mentally ill patients being discharged from hospital, and the Audit Commission's most optimistic day surgery targets would increase the total number of consultations each week by 5.6%: from 115 to 121.5 consultations. However, most of these consultations would be accounted for by shared care for asthma and diabetes, which does not represent extra workload. Patients were cared for in general practice before the 1990 general practitioner contract, and there would have been consultations with these patients whether or not there was a protocol for shared care. The effect of changes in day surgery and closure of long stay hospitals is an increase in the number of weekly consultations by only 0.8%—less than one extra consultation per general practitioner a week.

More evidence is still needed

The studies we reviewed cover only some of the changes that have happened in the 1990s, and the scope for making generalisations is limited. The studies evaluated schemes with well defined protocols and well resourced and organised facilities. This will not always be the case, and these are not the schemes that general practitioners are most worried about. Evidence is still lacking on the impact of early discharge from hospital, of day case surgery in which there is little regard to discharge arrangements, and of general practitioners taking over care of patients who otherwise might have been cared for in hospital—including visits to those in nursing and residential homes. Elderly people living in nursing homes require greater general practitioner input than does the remaining practice population over the age of 65,[40] and disabled people living in purpose built housing units increase general practitioners' workload more than do other newly registered patients.[41] These studies may indicate the workload implications of this type of work being transferred from hospital. However, studies are still needed which distinguish the work transferred from hospitals from that resulting from an aging population or patients moving house. Furthermore, in most studies, the definition of workload is confined to consultation rates, and other workload implications are not considered. Thus, very little is known about the extent of workload generated from the shifts in the balance of care.

Being unable to state whether a shift of services from secondary to primary care increases general practitioners' workload does not necessarily mean that there is not a problem. Rather, general practitioners have only anecdotal evidence to support their claims of appreciably increased workloads, and the NHS Executive has insufficient evidence to make informed decisions about remunerating general practitioners for the extra work resulting from the changes.

The present evidence does not point to significant increases in workload resulting from some of the most commonly implemented changes designed to shift the balance of care into general practice. More detailed analyses of the current concerns of general practitioners indicate that the secondary-primary care shift is not at the forefront of their thinking in relation to changed workload.[42 43] Other pressures, such as increased expectations of patients and administrative burdens, are of greater concern. As the shifts gain momentum the picture may change, and there will be an increased need to identify mechanisms for ensuring that resources follow workload.

The problem of remuneration

One way of addressing the issues arising from the movement of services into general practice, and the feeling of general practitioners that they are overworked, is to find ways of regulating the way in which general practitioners are paid, rather than simply paying them more. Identifying a level of compensation nationally is made difficult by the inevitable local variations in workload resulting from differences in how widespread the various schemes are, and by variations in the number of relevant patients on practices' lists. This problem would be partly solved by agreeing payment rates for providing non-core services. However, to be evidence based, this approach still requires more research on the work required in connection with specific schemes. Furthermore, this way of remunerating general practitioners would not take into account the use that general practitioners already make of skills within the primary care team to relieve themselves of some of their work, as recommended in *Primary Care: the Future*.[10] Tasks may be delegated to other professionals within the practice, who are already paid[44], freeing general practitioners to undertake new activities without concomitant increases in workload. Furthermore, greater use of "fee for service" remuneration is widely believed to lead to demand on these services induced by doctors as a means of increasing income. There is a danger of increasing health service costs with little impact on the health of patients.[45] Few studies have considered this assertion of perverse incentives empirically, and the results of those that have are questionable, due to their design and lack of controls.[46] There is, however, strong evidence from Denmark that a change from capitation to mixed capitation and fee for service does lead to increased service intensity.[47] The effects on use of resources and welfare of patients have not been examined empirically.[45 46 48] More research is needed to evaluate whether the effects of such a change in remuneration systems are compatible with the goal of value for money and effectiveness in health policy.

Delegation

An alternative way of looking at the issue of services moving into general practice is to increase the emphasis on delegation of work to other professionals. Aside from preventive work and chronic disease management, delegation to nurses or the employment of new professionals to take on shifted tasks is in its infancy. Moreover, the degree to which general practitioners delegate work varies across Britain. This approach may not reduce health service costs, but it should reduce the impact of a primary care led NHS on general practitioners' workload. It remains to be seen whether other structural changes within primary care can be harnessed to facilitate this process. What is clear is that additional audit and research are required so that claims are based on evidence and not anecdote.

"The impact of the secondary-primary care shift on GP workload" is available from the National Primary Care Research and Development Centre, 5th Floor,

Williamson Building, Oxford Road, Manchester M13 9PL.

Funding: No additional funding.
Conflict of interest: None.

1 NHS Executive. *Developing NHS purchasing and GP fundholding: towards a primary care-led NHS.* Leeds: Department of Health, 1994. (EL (94)79.)
2 NHS Executive. *Developing NHS purchasing and GP fundholding: towards a primary care-led NHS.* Leeds: Department of Health, 1995.
3 NHS Executive. *Towards a primary care-led NHS—briefing pack.* Leeds: NHS Executive, 1995.
4 Department of Health and Welsh Office. *General practice in the NHS: a new contract.* London: DHSS, 1989.
5 NHS Executive, Department of Health. *Practice budgets for general medical practitioners. Working paper 3.* London: HMSO, 1989.
6 NHS Management Executive. *GP fundholding practices: The provision of secondary care.* London: NHSME, 1993. (HSG (93)14.)
7 Secretaries of State for Health, Social Security, Wales, and Scotland. *Caring for people: community care in the next decade and beyond.* London: HMSO, 1989. (Cm 849.)
8 Secretaries of State for Health, Wales, Northern Ireland, and Scotland. *Working for patients.* London: HMSO, 1989. (Cm 555.)
9 NHS Executive. *A national framework for the provision of secondary care within general practice.* Leeds: NHSE, 1996. (HSG(96)31.)
10 NHS Executive. *Primary care: the future.* Leeds: Department of Health, 1996.
11 Secretary of State for Health. *Choice and opportunity. Primary care: the future.* London: Stationery Office, October 1996.
12 Petchey R, Williams J, Baker M. *Junior doctors, medical careers and general practice.* Nottingham Primary Care Research Unit, Department of General Practice, University of Nottingham, 1996.
13 Marsh GN. Flourishing or floundering in the 1990s. *Br J Gen Pract* 1992;42:266-7.
14 General Medical Services Committee. *Medical Workforce Task Group Report.* London: BMA, 1996.
15 Bower H. Dorrell prescribes more work for GPs. *Doctor* 1996 January 25:1.
16 DoH "improvements" to raise GP workload. *GP* 1996 January 19:1.
17 Elliot A. Resources should match vision, say GPs. *GP* 1996 January 26:2-3.
18 Warry R. GPs tell Malone they can't do more. *GP* 1996 February 23:2.
19 Husain O. "Dumping" fears are starting to hit home. *Doctor* 1996 February 5:84.
20 Lipley N. Few GPs welcome A&E work. *GP* 1996 February 9:2.
21 GPs fear being swamped with non-GMS work. *Pulse* 1996 June 15:7.
22 General Medical Services Committee. *Core services: taking the initiative.* London: BMA, 1996.
23 Pill RM, Stott NCH. *An investigation of the impact on GP and community services of increased day surgery.* Department of General Practice, University of Wales School of Medicine, 1995.
24 Diabetes Integrated Care Evaluation Team. Integrated care for diabetes: a clinical, social and economic evaluation. *BMJ* 1994;308:1208-12.
25 Grampian Asthma Study of Integrated Care. Integrated care for asthma: a clinical, social and economic evaluation. *BMJ* 1994;305:559-604.
26 Florey C du V, Yule B, Fogg A, Napier A, Orbell S, Cuschieri A. A randomised trial of immediate discharge of surgical patients to general practice. *J Public Health Med.* 1994;16:455-65.
27 Knapp M, Beecham J, Koutsogeogopoulu V, Hallam A, Fenyo A, Marks M, *et al.* Service use and costs of home based versus hospital based care for people with serious mental illness. *Br J Psychiatry* 1993;163:195-203.
28 Burns T, Raftery J, Bedsmoore A, McGigan S, Dickson M. A randomised trial of home based acute psychiatric services. II. Treatment pattern and costs. *Br J Psychiatry* 1993;163:55-61.
29 Burns T, Raftery J. Cost of schizophrenia in a randomised trial of home based treatment. *Schizophrenia Bulletin* 1991;17:407-10.
30 Dick P, Cameron L, Cohen D, Barlow M, Ince A. Day and full time psychiatric treatment: a controlled comparison. *Br J Psychiatry* 1985;147:246-50.
31 Hallam A, Knapp M, Beecham J, FenyoA. Eight years of psychiatric reprovision: an economic evaluation. In: Knapp M, ed. *The Economic evaluation of mental health.* Aldershot: Ashgate Publishing, 1995.
32 Beecham J, Knapp M, Allen C. Comparative efficiency and equality in community based care. In: Knapp M,ed. *The economic evaluation of mental health.* Aldershot: Ashgate Publishing, 1995.
33 Knapp M, Cambridge P, Thomson C, Beecham B, Allen C, Darton R. Residential care as an alternative to long stay hospital: a cost effectiveness evaluation of two pilot projects. *Int J Geriatric Psychiatry* 1994;9:297-365.
34 O'Cathain A. Evaluation of a hospital at home scheme for the early discharge of patients with fractured neck of femur. *J Public Health Med* 1994;16:205-10.
35 Audit Commission. *All in a day's work: an audit of day surgery in England and Wales.* London: HMSO, 1992. (NHS occasional paper No 4.)
36 Boyce J, Ralphs D, Devlin B. Day case surgery and workload. *BMJ* 1992;304:1174.
37 Fletcher J, Dawes M, McWilliam J, Millar J, Griffiths S. Day surgery and community health services workload: a descriptive study. *Br J Gen Pract* 1996;46:477-8.
38 Ghosh S, Sallam S. Patient satisfaction and postoperative demands on hospital and community services after day surgery. *Br J Surg* 1994; 81:1635-8.
39 Fry J. *General practice. The facts.* National Associating of Health Authorities and Trusts. Oxford: Radcliffe, 1993.
40 Andrew RA. Analysis of a general practitioner's work in a private nursing home for the elderly. *J R Coll Gen Pract* 1988;38:546-8.
41 Eyre S. An analysis of general practitioner's workload in a disabled housing development. *Br J Gen Pract* 1996;46:739-40.
42 Hayter P, Peckham S, Robinson R. *Morale in general practice.* Southampton: University of Southampton Institute for Health Policy Studies, 1996.
43 44 Evans D. *Changing work practices and workload in secondary and primary care. Report of a multi-agency working group in Southampton & South West Hampshire Health Commission.* Southampton: Institute for Health Policy Studies, University of Southampton, 1995.
44 Lilley R. Doctoring the system. *Guardian* 1996 November 13:2.
45 Hudhes D. General practitioners and the new contract: promoting better health through financial incentives. *Health Policy* 1993;25:39-50.
46 Scott A, Hall J. Evaluating the effect of GP remuneration: problems and prospects. *Health Policy* 1995;31:183-95.
47 Krasnik A, Groenewegen PP, Pedersen PA, van Scolten P, Mooney G, Gottschau A, *et al.* Changing remuneration systems: effects on activity in general practice. *BMJ* 1990;300:1698-701.
48 Langham S, Gilliam S, Thorogood M. The carrot, the stick and the general practitioner: how have changes in financial incentives affected health promotion activity in general practice? *Br J Gen Pract* 1995;45:665-8.

Socioeconomic determinants of health

Community marginalisation and the diffusion of disease and disorder in the United States

Rodrick Wallace, Deborah Wallace

Abstract

This article describes the cascading diffusion of "inner city problems" of disease and disorder in the United States—from the huge marginalised inner city communities of the largest municipalities, first along national travel routes to smaller cities, and then from central cities into surrounding more affluent suburbs—following the pattern of the daily journey to work. Public policies and economic practices which increase marginalisation act to damage the "weak ties" of the community social networks which bind central city neighbourhoods into functioning units. Spreading disease and disorder can be interpreted as indices of the resulting social disintegration, which is driven by policy. This "failure of containment" in the United States should serve as a warning for cities in Europe against reducing the municipal and other services that they provide to "unpopular" subpopulations.

This is the sixth in a series of eight articles edited by Richard Wilkinson

Introduction

Visitors from Europe who, by chance or design, encounter the marginalised poor communities of the

BMJ 1997;314:1341–5

continued over

New York Psychiatric Institute, New York, NY 10032, USA

Rodrick Wallace, *research scientist*

Public Interest Scientific Consulting Service Inc, New York, NY 10027, USA

Deborah Wallace, *president*

Correspondence to: Dr R Wallace, PISCS Inc, 549 W 123 St, Suite 16F, New York, NY 10027, USA. rdwall@ix.netcom.com

IMAGES

Commuting patterns predict the spread of diseases outward from inner cities to suburbs

large American cities are often stunned by the extent and intensity of physical and social deterioration. In New York City alone, some 600 000 people live in the devastated zones [1]. This "bombed out" urban landscape is largely the creation of public policies of "planned shrinkage" and "benign neglect," which led to reductions of services such as fire control and garbage collection in poor neighbourhoods.[2] Following similar service cuts and related disinvestment, much rental housing was rapidly destroyed during the 1970s in many large American cities. Presently, nearly a quarter of American children are, according to official figures, growing up in poverty, many within these collapsing communities.[3]

A principal underlying cultural assumption of the policies creating these conditions is that the effects of this devastation are being, and will be, confined largely to the targeted communities, and thus they will be separated from the suburban counties in which most affluent people now live and in which political power now lies. Only a quarter of Americans live in central cities; half now live in the surrounding suburbs.

A second underlying assumption is that, even if there is suburban "leakage" from the decaying central city neighbourhoods of New York, Los Angeles, Chicago, and some other large cities, this will not greatly affect the nation as a whole.

These cultural assumptions were epitomised by a study titled "The social impact of AIDS in the United States," published in 1993 by the National Research Council. The study concluded, on the basis of a single cross sectional map of AIDS in postal zones of New York city, that

> Many geographical areas and strata of the [American] population are virtually untouched by the epidemic and probably never will be [touched]; certain confined areas and populations have been devastated and are likely to continue to be .... HIV/AIDS will "disappear," not because, like smallpox, it has been eliminated, but because those who continue to be affected by it are ... beyond the sight and attention of the majority population.[4]

Geographic diffusion of disease and disorder

Such a statement flies in the face of a century of studies of geographic diffusion on several scales of space, time, and population, well summarised by Abler *et al*.[5] Almost needless to say, no geographers or spatial ecologists participated in the National Research Council's study. Geography, history, economics, anthropology, ecology, sociology, and epidemiology all study how rumours, fads, and technical and social innovations—as well as epidemics—spread in space and time and between social groups. Three mechanisms, acting at different scales, have been found to characterise such spread: hierarchical diffusion, spatial contagion, and network diffusion.

Hierarchical diffusion describes a cascading hopscotch transmission from socially dominant larger cities to smaller ones along the national transportation network. Two places may be geographically distant but they will be "close" in their probability of interaction if many people travel between them frequently. On a slightly smaller scale, spatial contagion (or expansion diffusion) describes radial spread along local travel routes from a central city epicentre into adjacent communities; this is often described as "a wine stain on a tablecloth." Network diffusion usually occurs on a still smaller scale; the term describes spread along personal, domestic, and community social nets which, when they have a geographic focus, can be characterised as sociogeographic structures.

Our recent work has quantified the spread of AIDS between the standard metropolitan statistical areas of the 25 most populous American cities, containing a total of 113 million people.[6] We found a hierarchical structure for the national AIDS epidemic: a top-down pattern of spread from the initially infected epicentres of New York City and San Francisco to other urban regions. Using data from the US Census Bureau on migration between metropolitan areas, we calculated the probability of contact between each region and all the others. The cumulative number of people with AIDS through 1995 within these metropolitan regions was closely predicted ($r^2 = 94\%$, the percentage of total variance in AIDS cases predicted through regression) by a model based on three logarithmically transformed variables: probability of contact with New York; probability of contact with San Francisco; and regional 1991 rate of violent crime per unit population. The spread of AIDS among the 25 largest metropolitan areas was thus determined by the intersection of local social disintegration (indexed by violent crime) and the probability of contact with the two most heavily infected epicentres.

The national and regional scales are bridged by the two determinants of the epidemic's structure: the links between regions and the socioeconomic structure, function, and history of the individual regions. Thus,

contrary to cultural assumption, large metropolitan regions with high prevalence of urban decay, such as New York, constitute great epicentres from which disease and disorder spread nationally.

When we examined population rates of AIDS, violent crime, and tuberculosis for the 24 counties constituting the New York City standard metropolitan statistical area, we found that a single composite index which convolved the area density of the workday commuting pattern with the local county poverty rate predicted well over 90% of the variance for each of these variables.[7] The intensity of the commuting pattern was determined from census data at county level on the daily journey to work; a step by step interaction was allowed to continue until a 24 element "equilibrium distribution" was reached.[7 8]

Breakdown of the AIDS data by time period (before 1985, 1985-7, 1988-90) showed that as the incidence of AIDS in Manhattan, the commuting center, rose, the incidence in all other centres rose correspondingly and in proportion (on a log-log scale, indicating a power law[7 8]). Figure 1 shows county rates of AIDS cases as a function of distance from Fifth Avenue and 42nd Street for 1982 and 1984, illustrating the nature of the relation, and figure 2 maps the spread of AIDS from the travel centre into those suburbs. Analysis for tuberculosis (1985-7, 1988-92) gave a similar result: as the incidence of AIDS in the dominant travel centre, Manhattan, rose, so in exact proportion did incidence in the counties in the entire metropolitan region.[7 8] The incidence of low birth weight, although not parallel to the three other markers, also was strongly predicted ($r^2 = 92\%$) by commuting pattern and poverty rate.

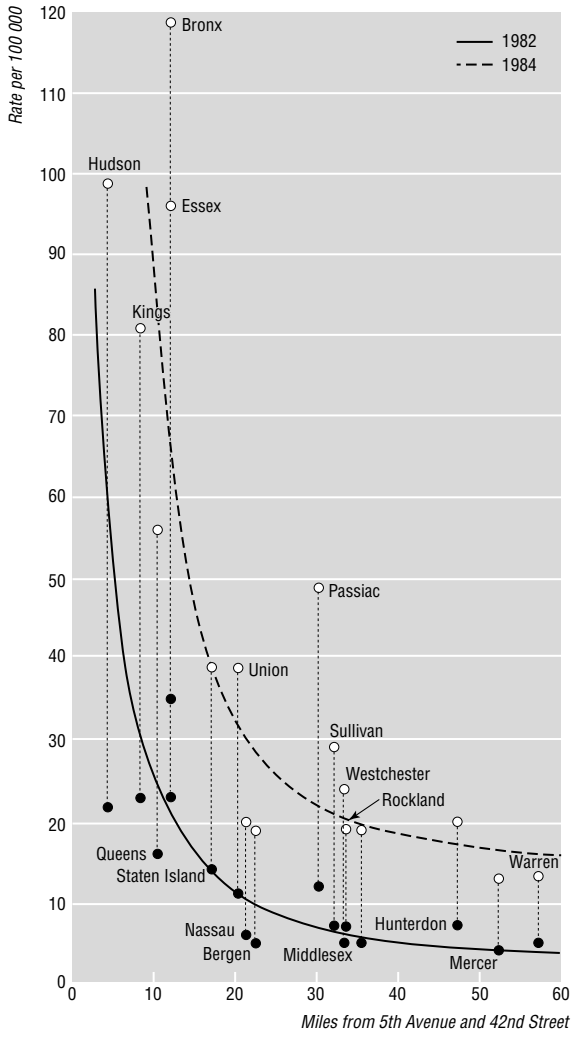

Fig 1 Cases of AIDS per 100 000 population for counties in or near New York City as a function of their distance from Fifth Ave and 42nd Street in Manhattan, 1982 and 1984. Rates for Manhattan rose from 131/100 000 in 1982 to 348/100 000 in 1984. Reprinted courtesy of Professor Peter Gould, Pennsylvania State University

Using the same approach, we analysed eight large standard metropolitan statistical areas in the United States, together containing 54 million people, for four public health problems: patterns of incidence of AIDS and tuberculosis, low weight birth babies per 10 000 live births, and the incidence of violent crime.[7] We characterised the public health problem of the central city as regionalised throughout the area if the statistical significance of the correlation data was not destroyed by removing the point of the commuting centre from the regression of the log incidence on the log of the commuting density per unit area. This omission of the travel centre constitutes a more rigorous condition, since such centres strongly dominate overall pattern. Regionalisation means that incidence in the central city determines the incidence in the surrounding counties, as modulated by the area density of the commuting pattern.

Different patterns of regionalisation emerged from our analysis of the eight areas.[7] Although all four public health problems showed regionalisation in the New York area, the other problems were regionalised only in some areas. At the other end of the spectrum from New York, the San Francisco area was regionalised only for tuberculosis.

Some areas showed trends toward regionalisation which indicate a strong influence of the central city on the outlying counties—for example, violent crime in the Washington DC and St Louis areas, and low weight births in the Detroit area. Some of these metropolitan areas are also characterised by a central city that is small in comparison to the total suburban population. Essentially, the travel centre tail wags the regional dog: disease rates in the core city and the local pockets of poverty in the county determine disease rates in suburban counties via the economic linkages within the region as indexed by the commuting pattern. Indeed, the workplaces of the metropolitan area mix the diverse populations. The area's single socioeconomic system is the reality: that city and suburb are totally separated is a public health myth.

Urban decay, social networks, and diffusion

Lives of individuals and families are deeply affected by influences on a neighbourhood scale: the neighbourhood embodies the "weak ties" through which the larger society channels information, support, and social control to families and individuals. These are relations of occupation, common interest, and neighbourliness beyond the "strong" ties of kinship, ethnicity, or peer group which bind small groups tightly and exclusively together into isolated "equivalence classes." Strong ties cannot easily serve larger

community purposes, a paradox which Granovetter characterises as the "strength of weak ties."[9]

As a neighbourhood disintegrates under the assaults of public policies of planned shrinkage and benign neglect, those weak ties begin to fray. Families leave, people are afraid to congregate on the streets, and legitimate economic activity (and the fraternity of occupation which it embodies) declines. As weak ties erode, possibilities for individuals and families narrow, and family groups are thrown back on their own resources.

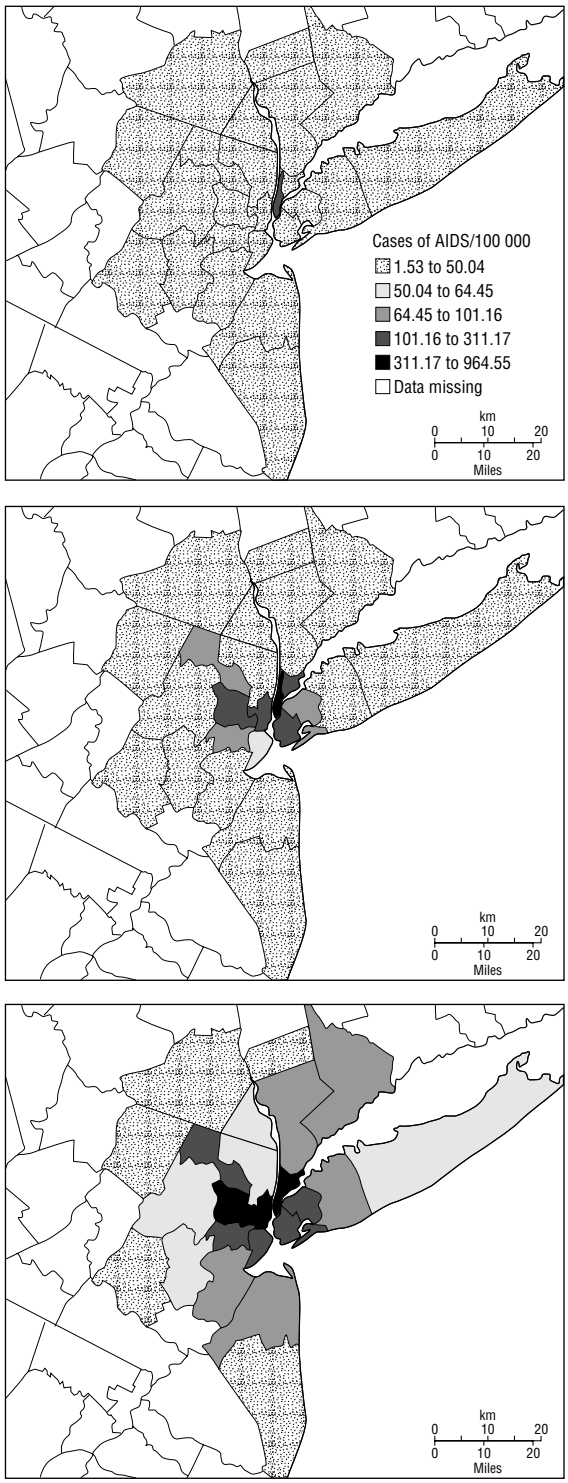

Fig 2 Cases of AIDS per 100 000 population by county for the New York metropolitan region (*a*) to end 1984; (*b*) 1985-7, (*c*) 1988-90

Youth behaviours such as doing well in school, getting a regular job, avoiding substance abuse, and maintaining stable relationships become more difficult as the neighbourhood structures that value such attainments dissolve. Negative acts such as violent behaviour, multiple sexual conquests, and drug taking are messages that can be more easily "heard" in a dissolving community than positive acts. If such "bad" behaviours damage a community's weak ties further—for example, by making street life more dangerous—the result may be destabilising positive feedback between community disintegration and antisocial behaviour.[10]

Neighbourhood processes affect families and individuals. Individuals and families who would otherwise have retained their housing become homeless due to the combination of the housing losses and the fragmentation of social networks. Individuals and families who would otherwise have remained independent and off the welfare roles have to receive public assistance as a result of a lack of low income housing and the disruption of community. Children who would otherwise have had one parent, if not two, become orphans from the epidemics of violence, substance abuse, and AIDS.

We find that public health at every scale of population is largely driven by contagious phenomena affecting socioeconomic processes, disease patterns, and behavioural processes at the neighbourhood level. Poor neighbourhoods in large central cities, suffering greatly from urban decay triggered and sustained by policy, have a disproportionate influence on the health, safety, and wellbeing of a huge proportion of the American population, including rich people.

With approaches from geography, demography, and ecology, the geographical patterns of disease can be modelled and predicted at a variety of scales of spatial distance, population, and socioeconomic distance. Disease and behavioural relationships between populations, whether purely spatially distant or merely socioeconomically distant, can be established.

In his seminal paper the noted ecologist CS Holling [11] described how, at each scale of space, time, and population, certain processes are crucial for stability and resilience. These processes are nested and linked between scales. In particular, Holling called attention to the "mesoscale," the familiar realm of population and community from a few metres to a few kilometres: at this scale, contagious processes funnel the impacts of events at the level of the individual and small group up to larger scales and also mediate events downward from the large scale to the micro, the individual.

Our data and analyses show that human ecology also includes fundamental processes, each of which have characteristic scales of population, geography, and time, and which are also nested and linked. As in natural ecosystems, contagious processes occur at the mesoscales of neighbourhood and city—one to a few kilometres—magnifying the impacts from the small to large and mediating impacts from the large down to the micro, the individual and the family.

In the United States, the keystone population which determines public health and public order at larger and smaller scales is the poor urban neighbourhood. If this structure cannot, for reasons of public policy and private interests, engage in the keystone

community processes, all populations in the country suffer deterioration of health and safety.

The belief that subpopulations in one country are separate and do not operate as a single ecosystem, affecting each other, has propelled the United States into a crisis of social and economic structure and of public health and public order which is so severe that even such crude measures as life expectancy show deterioration.[12] It reflects a profound error: concentration is mistaken for containment. Fundamental processes at and across the mesoscale ensure that concentration causes diffusion. Public policies or economic practices which marginalise vulnerable communities within Europe may be expected to create a crisis similar to that now raging in the United States.

The term "failure of containment" as applied to our analysis of the spread of social disintegration from poor urban neighbourhoods was first coined by Gregory Pappas.

Funding: The writing of this paper was supported under an Investigator Award in Health Policy Research from the Robert Wood Johnson Foundation.

Conflict of interest: None.

1 McCord C, Freeman H. Excess mortality in Harlem. *N Engl J Med* 1990;322:173-91.
2 Wallace R. Urban desertification, public health and public order: "planned shrinkage," violent death, substance abuse and AIDS in the Bronx. *Soc Sci Med* 1990;31:801-13.
3 Columbia University Center for Children in Poverty. *One in four: America's youngest poor.* New York: Columbia University, 1996.
4 National Research Council. *The social impact of AIDS in the United States.* Washington, DC: National Academy of Sciences, 1993.
5 Abler R, Adams J, Gould P. *Spatial organization: the geographer's view of the world.* Englewood Cliffs, NJ: Prentice-Hall, 1971.
6 Wallace R, Huang Y, Gould P, Wallace D. The hierarchical diffusion of AIDS and violent crime among US metropolitan regions: innercity decay, stochastic resonance and reversal of the mortality transition. *Soc Sci Med* 1997;44:935-47.
7 Wallace R, Wallace D, Andrews H. AIDS, tuberculosis, violent crime and low birthweight in eight US metropolitan areas: public policy, stochastic resonance, and the regional diffusion of inner-city markers. *Environment and Planning A* 1997;29:525-55.
8 Wallace R, Wallace D, Andrews H. Fullilove R. Fullilove M. The spatiodtemporal dynamics of AIDS and TB in the New York metropolitan region from a sociogeographic perspective: understanding the linkages of central city and suburbs. *Environment and Planning A* 1995;27:1085-108.
9 Granovetter M. The strength of weak ties. *Am J Soc* 1973;78:1360-80.
10 Wallace R, Fullilove MT, Flisher A. AIDS, violence and behavioral coding: information theory, risk behavior and dynamic process on core-group sociogeographic networks. *Soc Sci Med* 1996;43:339-52.
11 Holling C. Cross-scale morphology, geometry and dynamics of ecosystems. *Ecolog Monogr* 1992;62:447-93.
12 Wallace D. Smaller increases in life expectancy for blacks and whites between the 1970's and 1980's. *Am J Public Health* 1990;85:875-6.

Any questions

Why does rebound abdominal pain occur after anti-ulcer drugs are stopped?

Our attempts to stop proton pump inhibitors and H_2 antagonists in patients with dyspepsia (but not proved ulcers) are often associated with a rebound in abdominal pain. What are the likely causes of this? Are high gastrin concentrations (resulting from suppression of hydrochloric acid) likely to be important?

The causes of non-ulcer dyspepsia are numerous, but the commonest cause associated with rapid relapse of symptoms is gastro-oesophageal reflux. Clinically, the epigastric pain is usually associated with symptoms of heartburn, regurgitation of acid, and increased oral flatulence. Such symptoms may occur without the endoscopic changes of oesophagitis or hiatus hernia, but the diagnosis may then be confirmed by a correlation between symptoms and a fall in pH on monitoring of oesophageal pH. Symptoms frequently recur one to seven days after proton pump inhibitors or H_2 receptor drugs are stopped and are due to a recovery in secretion of gastric acid resulting in recurrent acid reflux. In such patients continuous medical treatment is usually needed long term or open laparoscopic antireflux surgery may be considered.

Areas of endoscopic gastritis or duodenitis may sometimes be associated with epigastric pain. As with peptic ulcers, the relapse of symptoms after treatment is stopped often takes several months to occur. Other causes of dyspepsia include functional pain; doctors should ask patients whether their treatment is having a definite effect and also consider whether there could be a placebo response.

Gastrin is secreted by the G cell of the gastric antrum and to a lesser extent from the duodenum. It is one of several chemicals that stimulate secretion of gastric acid; others include acetylcholine through vagal cholinergic neurones, γ-aminobutyric acid, histamine, and thyrotrophin releasing hormone. Gastrin has numerous other effects, including the stimulation of secretion of water and electrolytes from the stomach, upper small intestine, and pancreas; the inhibition of the absorption of water and electrolytes from the lower small intestine; the contraction of the lower oesophageal sphincter; and the stimulation of mucosal growth.[1]

Concerns have been expressed that the increased concentrations of gastrin that are associated with inhibition of gastric acid by modern drugs might result in acid rebound or be associated with carcinogenesis or cause more rapid growth of tumours. For example, omeprazole causes an important but modest rise in fasting gastrin concentration, which returns to normal eight weeks after the drug has been stopped,[2] and stopping the drug is not associated with an overswing in secretion of acid.[3] Much higher concentrations of gastrin are seen in patients with gastrinomas or pernicious anaemia. The modest rises in gastrin that occur with acid blockade are not thought to have adverse effects such as acid rebound (except rarely in patients with a gastrinoma) or have an important role in carcinogenesis.

R G Long *is a consultant gastroenterologist in Nottingham*

1 Bryant MG, Adrian TE. Gastrin. In: Bloom SR, Long RG, eds. *Radioimmunoassay of gut regulatory peptides.* London: W B Saunders, 1982: 36-41.
2 Sharma BK, Walt RP, Pounder RE, Gomes M de FA, Wood EC, Logan LH. Optimal dose of oral omeprazole for maximal 24 hour decrease of intragastric acidity. *Gut* 1984;25:957-64.
3 Prewett EJ, Hudson M, Nwokolo CU, Sawyerr AFM, Pounder RE. Nocturnal intragastric acidity during and after a period of dosing with either ranitidine or omeprazole. *Gastroenterology* 1991;100:873-7.