# Changes in medical practice following superficial and deep processing of evidence:

# a controlled experiment in clinical guideline implementation.

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of M.D. by Carl William Reginald Onion.

May 1997

*Abstract:*

**Changes in medical practice following superficial and deep processing: a controlled experiment in clinical guideline implementation**

*Carl WR Onion. M.D. thesis. December 1995. University of Liverpool.*

**Introduction**

Medical practice responds slowly to knowledge arising from medical research. Continuing medical education, and in particular clinical guidelines, are viewed as means to enable scientific advances to influence general practitioners' (G.P.) practices.

Medical educators are losing faith in traditional (superficial processing) teaching techniques which rely upon transmission of facts. Interactive methods, which involve students in deep and critical appraisal of facts (deep processing), are becoming more popular.

Psychology research suggests that deep processing might result in more favourable attitudes and consistent, persistent behaviour change.

**Method**

All 69 general practices in a district were randomly assigned to either a control (no intervention) group, a 'superficial' group, or a 'deep' group. Superficial practices were sent printed material and were invited to attend a lecture event lasting 1 hour. Deep practices were sent printed material and were offered a one hour discussion in their own practices with medical educators.

Verbal evidence of depth of processing was assessed. Attitudes regarding management of common infections was measured by questionnaire prior to intervention and 12 months after. Relevant prescribing and laboratory investigation data was collected in the year prior and year following intervention. A clinical study of cystitis recorded treatment failure rates.

**Results**

One third of superficial group GPs attended the lecture; all deep group practices were visited. There was evidence of deeper processing in the deep group. Knowledge and attitudes to infection treatment showed clearly favourable changes. Advocated antibiotics became more commonly prescribed, disapproved antibiotics less; the effect still evident aftert 9 months. Cystitis patients of the deep group experienced significantly less treatment failures. Prescribing cost was less than expected in the deep group.

There were no significant changes in knowledge or attitudes following superficial processing. Change in prescribing of all advocated relevant antibiotics were more favourable in the deep group except in the case of co-amoxiclav. Disapproved prescribing was not affected by superficial processing, neither was laboratory investigation or cystitis patient outcome. There were no prescribing cost savings.

**Analysis**

Deep processing was effective at achieving persistent attitude change with consistent alterations in prescribing behaviour and improved patient outcome. Superficial processing was associated with a brief novelty effect, but no evidence of attitude change and no improvement in patient outcome.

**Discussion**

If good medicine is marked by adoption of evidence based therapies, appropriate abandonment of superseded treatments and caution in taking up new ones then the deep group were better doctors for their intervention, the superficial were not.

Medical educational strategy, and clinical guideline implementation, must incorporate deep processing of issues by recipients to be effective.

# Contents

# Acknowledgements

The research and writing of the thesis, the design and execution of the experiment, and the interpretation of the data are my own work. However, the endeavour would have been impossible without others' co-operation. I am grateful to the following for generously contributing their wisdom, expertise, and time.

**Dr Debra Ashby** (Statistician, Liverpool University), **Prof. John Bligh** (Professor of Medical Education, Liverpool University), **Ms Collette Brooks** (Research Psychologist, Liverpool University), **Ms Darys Broomfield** (Research Psychologist, Liverpool University), **Dr Iain Buchan** (Statistician and Clinical Fellow, Liverpool University), **Mr Frank Burns** (Chief Executive Officer, Wirral Hospital Trust), **Sir Cyril A Clarke** (Past President, The Royal College of Physicians), **Dr Frananne Damato** (Medical Audit Facilitator, Wirral Health), **Mr Richard Donegan** (Systems Development Manager, Prescription Pricing Authority), **Miss Clare Dutton** (Pharmaceutical Adviser, Wirral Health), **Dr A Gordon England** (Regional Scientific Officer, North-West Regional Health Authority), **Mr Phil Evans** (Research Educationalist, Liverpool University), **Mr David Fillingham** (Chief Executive Officer, St Helens and Knowsley Health Authorities), **Dr Keith Foggin** (General Practice Clinical Tutor, Wirral), **Dr Trevor Gibbs** (Associate Postgraduate Adviser in General Practice, Liverpool University), **Prof. Conrad Harris** (Professor of General Practice, Leeds University), **Ms Melanie Iredale** (Research and Development Manager, Wirral Health), **Dr Jacqueline Jolleys** (Chairman, Association of Primary Care Medical Advisers), **Dr**

**Molly McBride** (Hon. Secretary, The Royal College of General Practitioners), **Mr Chris McCabe** (Health Economist, Sheffield University), **Mr Craig McIntosh** (Principal Scientist, Wirral Hospital Trust), **Dr Alan Murray** (Consultant Microbiologist, Wirral Hospital Trust), **Ms Sarah Peters** (Research Psychologist, Liverpool University), **Prof. Peter D Slade** (Professor of Clinical Psychology, Liverpool University), **Mr Mark Stead** (Chief Medical Laboratory Scientific Officer, Wirral Hospital Trust), **Mrs Angela Tinkler** (Research Assistant, Wirral Health), **Dr Diana Walford** (Director, Public Health Laboratory Service), **Prof. Tom Walley** (Professor of Clinical Pharmacology and Pharmaco-economics).

I am grateful to **Prof. RWI Cooke** and the Small Grants and Research Training Committee for supporting the study, and to **Prof. Tony Hart** for placing the facilities of the Department of Medical Microbiology at the University of Liverpool at my disposal.

Particular thanks are expressed to **Dr Christos Bartzokas,** colleague and mentor and to **Dr Godfrey W Smith** for his support and guidance.

Above all must be acknowledged the support and fortitude of **Gaynor,** and to **Jessica, George** and little **Penny** for tolerating a husband and father obsessed with 'deep processing' for six years to the exclusion of the nice 'superficial' things in life.

# Chapter 1

# Differences between science and practice, and between practitioners

> *Men act like brutes in so far as the sequences of their perceptions arise through the principle of memory only, like those empirical physicians who have practice without theory.*[1]
>
> Gotfried Leibniz (1714)

## Summary of chapter

Scientific evidence is slow to influence clinical practice, and practitioners respond at different rates. These variations are partly due to differences in the situations practitioners' work in, and partly to their individual dispositions. Situational factors require practical attention, dispositional require fundamental changes in individuals' attitudes. Attitudes are most reliably changed through intrinsic motivators, such as alignment of the desired outcome with the practitioners' personal professional goals. To achieve this would require a major shift in continuing medical education as most practitioners still prefer passive education to the interactive work that is most likely to alter attitudes. Deep processing (effortful thinking on the personal relevance and ramifications of an issue),

which is the key ingredient in interactive small group work and problem-based learning, is associated with better retention of information and with subsequent attitudes that are persistent, resistant to counter-persuasion, and are associated with consistent behaviour.

Superficial processing (passive receipt of information), which is the essence of lectures and reading, is associated with short-lived attitudes that are susceptible to counter-persuasion, and with inconsistent behaviour.

## 1.1 Variation in Medical Practice

Medical research has been defined as the discovery of the 'right' medicine for a condition, and medical audit as measuring whether the 'right' medicine is applied in practice.[2] If medical practitioners do not practise in accordance with scientific recommendations then medical practice is sub-optimal. The aims of continuing medical education are to inform practitioners of recent evidence and to persuade them to alter their practice accordingly. Sub-optimal practice will occur if current continuing medical education is insufficient to inform or persuade doctors to change their practices according to recent evidence from science. If this is the case, that medical science is not followed directly by concomitant practice, surely medical science represents a good deal of wasted effort. If there is discord between scientific recommendations and medical practice then, for patients' sake, the discovery of how to inform and persuade doctors to follow such recommendations more effectively becomes an urgent and necessary pursuit.

Variation is caused by differences amongst doctors over decisions on:

- whether or not to investigate?
- what investigation to carry out?
- whether or not to prescribe treatment?
- what treatment to prescribe?
- whether or not to refer?
- to whom to refer?

The following sections will explore variations in prescribing in particular to throw light on what might influence this decision-making because variation in prescribing in general practice has been highlighted as an area of concern[3] with great scope for improvement.[4]

## 1.2 Variations in prescribing.

Ideally, research clarifies the best treatment options for any given condition. Doctors quickly become aware of the current research evidence and adjust their practices accordingly. Consequently, the idealist might make the following logical assertions:

*Assertion I* - all clinicians prescribe the same defined range of drugs.

*Assertion II* - their prescribing changes to match current scientific recommendations.

In reality, the opposite occurs.

In the study below I found that striking variations in prescribing exist amongst general practitioners (GPs), even those in the same geographic area with similar patient demographics. This is both in terms of volume of prescribing and the type of drugs prescribed. Antibiotic prescribing is particularly variable.

To quantify variations in prescribing rate of different classes of antibiotics amongst general practices I undertook the following study.

Study details

**Antibiotic prescribing in general practice: variation among GPs in their prescribing rates for different classes of antibiotic.**

- Researcher: CWR Onion

- Population: 18 practices covering a population of 58,000. Half (9) of the practices were single-handed.

- Setting: Wallasey, a well defined cluster of small centres of population at the tip of the Wirral peninsula.

- Method: observational survey.

- Data: 'PACT' data from the national Prescribing Authority database of all dispensed NHS prescriptions issued by the GPs in the quarter ending April 1992. Data expressed as prescription items per 1000 population

- Analysis: Two practices did not prescribe quinolones in the study quarter, otherwise all practices prescribed all classes of drugs. Two hundred and thirty-four prescriptions per 1000 population in the antibiotic classes under study were prescribed by the GPs. Broad spectrum penicillins were most commonly prescribed, and quinolones least.

There was a considerable range in prescribing rate within each class (see table 1.2a). Single-handers only differed from multiple partnered practices in that they were more likely to be in the extreme centiles of the variation, this effect could be expected, from a statistical point of view, because their data would not be diluted by other doctors' prescribing.

Table 1.2a Spread of number of prescriptions per 1000 population among 18 practices in one locality in a single quarter (April 1992). Total antibiotic prescriptions = 13,463, total population 57,551.

| Antibiotic | Median | Lower quartile | Upper quartile |
|---|---|---|---|
| Broad spectrum penicillins | 81 | 65 | 98 |
| Macrolides | 25 | 16 | 44 |
| Cephalosporins | 22 | 12 | 57 |
| Tetracyclines | 20 | 12 | 39 |
| Penicillin V | 11 | 2 | 21 |
| Co-trimoxazole | 10 | 4 | 27 |
| Trimethoprim | 9 | 3 | 14 |
| Penicillinase-resis. penicillins | 6 | 2 | 10 |
| Quinolones | 5 | 1 | 12 |

•Conclusion: there is considerable variation amongst practices in the prescribing rate within all classes of antibiotics

If there is no consistent approach to antibiotic prescribing in general practice it might be that the prescribing of individual GPs changes at different rates, hence an apparently irrational variation.

However, it has been shown that individual GPs change their customary drugs in only 5.4% of new cases,[5].The authors of the study concluded that GPs' prescribing habits were 'stable and conservative'.* GP prescribing is not confined to a defined range of drugs, and does not change quickly to match scientific recommendations.

Prescribing diversity is said to be rational when the range of drugs are of equal merit with regard to efficacy, appropriateness, acceptability, safety and economy.[6] Demonstrating 'irrational' prescribing can be difficult because of a tendency for practitioners to rationalise their prescribing in retrospect. For instance pharmacologically inactive drugs, the cerebral vasodilators, are prescribed in substantial numbers at enormous cost in Scotland. An author justified this on the grounds of their *placebo effect* in peripheral vascular disease.[7] Use of such expensive preparations as placebos cannot be rational when there are less expensive, less noxious alternatives with convincing names such as 'tonics' or vitamins.

Trimethoprim is just as effective as co-trimoxazole (trimethoprim plus sulphamethoxazole) in urinary tract and respiratory tract infection with less side effects,[8] so one would

---

* See Appendix A for a critical appraisal of this paper

expect it to be prescribed much more frequently than co-trimoxazole. Co-trimoxazole has recently been prescribed as much as pure trimethoprim. In the Wirral district in 1992 GPs prescribed trimethoprim as 5.6%, and co-trimoxazole as 5.8%, of antibiotics (analysis by the author in 1992 - data source: Prescription Pricing Authority). Hospital practice is just as irrational. A study of the antibiotic treatments of patients with urinary tract infections was performed in a large teaching hospital in Liverpool in 1986. Half the infections were community acquired. The doctors prescribed 16.2% more co-trimoxazole than trimethoprim to no greater clinical advantage and at much increased cost in both financial terms and in terms of the risk of adverse effects[9]

"How should acute otitis media be treated?" was asked of 85 trainers in general practice in 1983; 98.8% of them said that an antibiotic was necessary; and 70.6% of them declared that penicillin V (phenoxymethylpenicillin) was their drug of first choice.[10] The antibiotic is only going to be effective against bacterial cases of which a common pathogen is *Haemophilus influenzae*. *H. influenzae* had long been known as highly resistant to penicillin,[11] so penicillin V would not seem to have been a well-informed choice. Ampicillin would have been more appropriate.

In 1993 the author performed a local survey attitudes to the best treatment of some common infections in the Wirral peninsula.

Study details

**Empirical treatment of common infections: GPs' preferred treatments**

•Researcher: CWR Onion

•Purpose: to establish the range and popularity of antibiotics among GPs as empirical treatment (i.e. without laboratory test results as guidance) for common infections.

•Subjects: All 179 GP principals in the Wirral district.

•Method: Open postal questionnaire survey. The questionnaire asked the GP to write his or her chosen first-line oral antibiotic next to a list of 8 common infections.

•Data: One hundred and forty-four responses were received (80.4% response rate). No exclusions were necessary.

•Analysis: the range (number {'range'} of different antibiotics mentioned) and 2 most preferred antibiotics (as proportions of responses) are shown in the table below. Other details in following text.

| Infection | Range | Most popular % | 2nd most popular % |
|---|---|---|---|
| Bronchopneumonia | 19 | Co-amoxiclav 46% | Amoxycillin 24% |
| Bronchitis | 17 | Amoxycillin 34% | Co-amoxiclav 32% |
| Lobar pneumonia | 17 | Co-amoxiclav 42% | Amoxycillin 16% |
| Sinusitis | 12 | Doxycycline 43% | Amoxycillin 17% |
| Acute otitis media | 10 | Amoxycillin 52% | Co-amoxiclav 20% |
| Cystitis | 9 | Trimethoprim 71% | Nitrofurantoin 6% |
| Cellulitis | 7 | Flucloxacillin 46% | Co-amoxiclav 13% |
| Tonsillitis | 6 | Penicillin V 76% | Amoxycillin 12% |

•Conclusion: GPs exhibit a wide range of preferred antibiotics for common infections though there are clear favourites. No inference can be drawn from this study about what is actually prescribed for these conditions.

In this study, 5% of general practitioners had asserted that penicillin V was their first choice antibiotic in acute otitis media. Scientific evidence is clearly slow to influence practice fully.

In the same study 15% of Wirral practitioners stated that their empirical antibiotic in tonsillitis was amoxycillin or ampicillin. Glandular fever is a frequent cause of tonsillitis in general practice and can be difficult to diagnose in its early stages. So, in spite of the warning in the British National Formulary of severe rashes caused by the use of amino-penicillins in glandular fever, 15% of the doctors exposed their tonsillitis patients to an unnecessary risk, an action that could have proved difficult to defend in court.

In lobar pneumonia 8% declared their first choice antibiotic as ciprofloxacin despite mounting evidence of treatment failures in

*pneumococcal* infection. Another example of science failing to influence practice, this time in the management of a commonly fatal condition.

Large variations in the management of common conditions are evident between practitioners, but are individual practitioners variable ? A recent Norwegian study used thoroughly prepared actors to present unexpectedly for consultation with 52 GPs (who had volunteered 1 year previously). A bogus patient visited each once and a second bogus patient with identical characteristics and presentation attended 6 weeks later. Their performances were scored against a clinical guideline. The consistency amongst the bogus patients and between the doctors as a group was high, but there was a significant intra-doctor difference between the consultations in the way the doctor managed the case, including the treatment prescribed.[12]

Some prescribing actions defy reason and yet contribute to the variation in prescribing patterns. The above demonstrates that a significant component of variation in clinical practice is irrational.

## 1.3 Situations and dispositions: the attribution theory.

The fact that there are variations in behaviour implies that there are differences in the *situations* of individual GPs or their *dispositions* or both; this is the axiom of the attribution theory of psychology.[13] Attitudes are dispositions, and it is necessary to change attitudes to successfully influence behaviour as we shall see later. "In essence, an *attitude* is an *evaluative disposition toward some object.* It's an evaluation of something or someone along a continuum of like-to-dislike or favourable-to-unfavourable."[14]

A regression analysis of situational factors and differences in prescribing rates and costs between different family practitioner committees (the district level aggregations of GPs) was performed in 1991.[15] It was discovered that 51% of the variation in prescription rates and 44% of the variation in costs per patient could be explained by the following factors:

1. Age or sex demographics
2. Standardised Mortality Ratio (all causes)
3. Number of GPs per head of population

but not by deprivation (Jarman index) scores. It would appear that about 50% of the variation in prescribing behaviour

could be due to individual GPs' dispositions. (See appendix A for a critical appraisal of this paper).

Attribution theory gives two useful pointers to how we might ascertain whether an individual GP's unique prescribing behaviour results from situation or disposition. *Non-normative* behaviour, or highly unusual behaviour, suggests a dispositional cause; the individual has a non-conformist disposition. For example, in one Wirral locality (in 1992) GPs prescribed a mean of 11.6 (standard deviation 21.6) quinolone prescriptions per 1000 population per quarter. The likelihood of a practitioner prescribing more than the mean plus three standard deviations (i.e. 76.4) is less than 1%. However, one practitioner prescribed 93.2 prescriptions per quarter. The attribution theory predicts that this is probably due to a dispositional factor. The evidence enabled the targeting of help to the doctor. It transpired he was suffering from emotional distress and early retirement on sickness grounds was facilitated for him. *Consistent* behaviour over time and in different situations also suggests that disposition is triumphing over situation. Behaviour is not changing in response to changing situations and, as noted above, GPs are known to be 'conservative and stable' prescribers. A third indicator of strong dispositional factors is *non-distinctive* behaviour, a similar response to distinctively different

situations. For example, one local practitioner has a high rate of prescribing, a high rate of prescribing new drugs, a high rate of requesting domicillary visits and a high rate of use of laboratory investigations. In this case examination of several different factors displays a fundamental attitude which affects his performance across the board. This is not likely to change with any outside efforts to alter his situation. An American study in 1987 found that the *rate* of change of prescribing behaviour was independent of most individual practitioner characteristics, including age, board certification, speciality, and location of practice.[16] In other words the rate of adoption of improved practice depends more on disposition (i.e. attitudes) than circumstances.

So, variations can be due to two classes of factors. Situational factors (which can be identified by a top-down district comparison approach), and dispositional factors (which are identified by an individualised approach). The major situational factors (such as area of practice, place of training, workload etc.) have been identified, but the American experience is that situational factors do not affect the rate of adoption of improved prescribing practices.

Attitudes are types of dispositions arising from evaluations of issues, as opposed to dispositions arising from personality type or

mental illness for example. Attitudes are the only dispositions amenable to continuing medical education and so are key issues in this thesis.

## 1.4 Attitudes to prescribing and clinical guidelines

'Attitude' is the judgement a person holds towards an issue or object. Prescribing is the commonest concrete intervention in general practice and so attitudes to this activity (rather than referral, patient education etc.) followed up with examination of guidelines in general will be described in this section.

A questionnaire survey by Taylor (published in 1981) of Aberdeen GPs divided respondents into four classes according to their prescribing characteristics and sought associated factors.[17] The classes were high volume and high cost per item; high volume and low cost; low volume and high cost; and low volume and low cost. A nested study examined the personal NHS prescribing of 14 of the experienced principals in December 1974 in more detail — volume was primarily calculated as number of prescription items per 1000 patients on the doctor's list and cost as the total ingredient cost per 1000 patients.[18] Each class had distinctive characteristics:

Table 1.4a Classification of general practices by rate and cost of prescribing with associated features.[15]

| Class | Quantity and Cost | Characteristics: |
|---|---|---|
| A | High volume/high cost | • Recently qualified (<10 years)<br>• Smaller partnerships<br>• Above average no. of patients |
| B | High volume/low cost | • Low proportion of new drugs<br>• Poor quality of practice scores |
| C | Low volume/high cost | • High proportion of new drugs<br>• Best quality of practice scores |
| D | Low volume/low cost | • Urban practices<br>• Teachers or Royal College members<br>• Low prescribing rate for antibiotics and cardiac drugs. |

The nested study identified that the therapeutic area with the greatest variation in cost identified was antibacterial drugs.

Group A are younger and busier, the throughput may drive the high rate, and recent hospital experience the high use of expensive drugs. Group B exhibit poor quality across the board (including practice administration and record keeping), they may be unaware of new treatments and alternative therapies, and prescribe drugs liberally. Group C appear to prescribe carefully and be aware of new drugs. Group D seem similar to group C in that they are careful prescribers, but consider cost and no-treatment options more. The characteristics in table 1.4a represent a mixture of situational and dispositional factors. How much of these differences are associated with differences in

attitudes to prescribing related issues, such as formularies or clinical guidelines ?

Study details

**General practitioners and clinical guidelines: a survey of prescribing and attitude systems**

- Researcher: CWR Onion
- Purpose: (a) to ascertain GPs' attitudes to clinical guidelines for each of 5 components of an attitude system and (b) to explore the relationship between overall prescribing characteristics and attitudes to guidelines.
- Subjects: Ninety randomly selected 90 GP principals in Wirral district in 1992.
- Method: structured questionnaire survey (5 questions with answers on 5 point scales from very positive to very negative) in 1992. More details of the design are presented in chapter 4.
- Data: Questionnaire responses and PACT data at practice level for cost per prescribing unit (individuals over 65 weighted x3, those under 65 weighted x1) and number of prescription items per prescribing unit (PU) during the 12 months prior to the survey. Individual GPs were ascribed the prescribing characteristics of the practice they were in.

- Analysis: A response rate of 87% (78) was achieved. Low volume prescribers had a significantly more positive attitudes towards guidelines overall (70%) than high volume prescribers (57%, chi squared test P=0.05). However, there was no difference between high and low cost prescribers. Individuals' prescribing attitudes were probably diluted by their partners' prescribing because the prescribing data was collected at practice level. More details on individual comments are presented in chapter 4.
- Conclusion: GPs with a lower rate of prescribing (than the local mean items per PU) have distinctly more positive attitudes towards clinical guidelines than those with a higher rate. However, there is no difference between GPs with lower prescribing cost (than the district mean cost per PU) and those of higher cost.

A survey by Weiss *et al* in 1996 found that high cost and high volume practices are substantially more concerned about financial constraints and incentives around prescribing (see next section1.5).[19] Weiss's study and my guideline attitude study above support Taylor's conclusion that there are distinct characteristics associated with GPs who are high cost or high volume prescribers relative to their colleagues.

## 1.5 Motivation.

In 1989 Pitts and Vincent performed a study on influences on prescribing for sore throats in general practice.[20] The authors' conclusion was that it is much easier to influence which drugs GPs prescribe than their *rate* of prescribing. Some outside influences appear to affect qualities of prescribing, but not rate of prescribing. The rate is related to attitude (disposition). It is more likely to be affected by *intrinsic* motivators. Motivation is said to be intrinsic when a person performs an act purely to manifest competence and self-determination. The application of an extrinsic motivator (such as financial reward) to induce someone to undertake a task that he is intrinsically motivated to perform can sometimes paradoxically reduce compliance or performance.[21]

However, a successful drug formulary implementation initiative in an English health district showed impressive favourable changes in the types of drugs prescribed *and* in the number of prescription items per head of population (up to 12%, depending on therapeutic area, in both cases).[22] Why the rate changed in this case is uncertain, but the presence of background financial incentives to reduce prescribing cost at that time (national fundholding and prescribing incentive schemes) may have played a part — financial incentives associated with

numbers of items prescribed have been shown to be potent reducers of prescribing rate in Medicaid primary care practitioners in the USA.[23]

In an English survey (59% response rate) by Weiss and colleagues of 386 GPs' perceptions of 4 broad areas of concern associated with prescribing; namely their

1. sense of burden in providing health care
2. views on financial constraints and incentives
3. use of prescriptions to control clinical workload
4. perception of demanding patients

only the respondents' concern about adverse financial pressures on prescribing decisions was related to actual prescribing behaviour — those concerned about these financial incentives and constraints prescribed less drugs by generic name, had higher practice prescribing costs per head of population and prescribed more items per head of population than the district average. The authors concluded that financial constraints and incentives around prescribing introduced in the previous 3 years had generated a significant amount of concern among GPs with unfavourable prescribing practices, but had not necessarily improved their actual behaviour.[24]

Evidence concerning the effect of financial incentives on prescribing rate is therefore equivocal.

A study by Mills and Chaffe (1993) in a large Sheffield hospital was designed to determine whether cost-consciousness was improving amongst anaesthetic practitioners following cost-awareness publicity in the department.[25] A survey of accuracy of estimation of cost of a 'shopping basket' of drugs and equipment was performed in 1987. It was repeated in 1993 and the results compared. There had been no improvement in cost awareness in that time (in fact costs were mainly over-estimated). A similar survey of physicians in Spain regarding their knowledge of the costs of they drugs they prescribe — there were 5 choices of price for each drug — showed only an accuracy rate of about 60%, the authors considered this level of knowledge to be deficient.[26] These studies suggest that doctors are not intrinsically motivated to learn the costs of drugs.

According to Ford, motivation regarding various tasks varies from person to person according to his own set of personal goals, the emotion aroused by those goals and how achievable he perceives the goals to be.[27] An accountant or storeman will have financial efficiency as a goal and will therefore be motivated to become familiar with the prices of goods, but what about doctors'

goals? Ford is currently refining an appropriate research tool (at Stanford University) — a manual of questionnaires to identify individuals' goals — when complete it will offer a valuable insight into the actual motivational goals of GPs, but for now these goals must be surmised by guesswork or, more satisfactorily, deduced from related types of study.

In a relevant study, Bradley in Manchester in 1992 invited GPs to undergo cathartic examination of recent uncomfortable prescribing decisions.[28] GPs would have felt most dissatisfied where progress towards their personal goals was not achieved. Seventy-four GPs recalled 307 critical incidents that made them feel uncomfortable. The reasons for concern were:

1. **Possible drug toxicity (27%)**
2. **Failure to live up to own expectations (21%)**
3. **Possible inappropriate medication (20%)**
4. **Ignorance / uncertainty about the management (17%).**

The commonest drug category causing concern was antibiotics (70% of all these critical incidents).* The areas listed can be expressed in terms of Ford's goals that we can assume to have been frustrated. These are:

---

* See appendix A for a critical appraisal of this paper.

1. **Social responsibility**
2. **Positive self-evaluations**
3. **Mastery (of the craft of medicine)**
4. **Self-determination (having the situation under control).**

This has been borne out by a local (currently unpublished) stress management initiative for GPs that I have been managing with colleagues. Several small groups of GPs discussed stressful aspects of their lives and possible solutions with experienced clinical psychologists. Recurring themes have been their feelings of being unable to match up to the social expectations of their role, finding difficulty in establishing whether they were good at their jobs or not - with consequent poor levels of self-esteem, and feeling helplessly swamped by increasing patient demands.* These expressions of frustration support the assumptions about GPs' goals drawn from Bradley's work.

They are laudable goals for a doctor, but not necessarily for a captain of industry. His goals might be superiority, resource acquisition, material gain, and individuality. Cost-consciousness is directly aligned with the goals of the businessman, but not with the doctor. Economy is therefore likely to be a poor motivator for a GP unless linked in some way to his goals. A persuasive argument for a GP might be, "You have a responsibility to society not waste

NHS money, and if you do make savings on prescribing you can spend these on patient care as you wish". This link is explicit in the GP fundholding scheme in the UK.

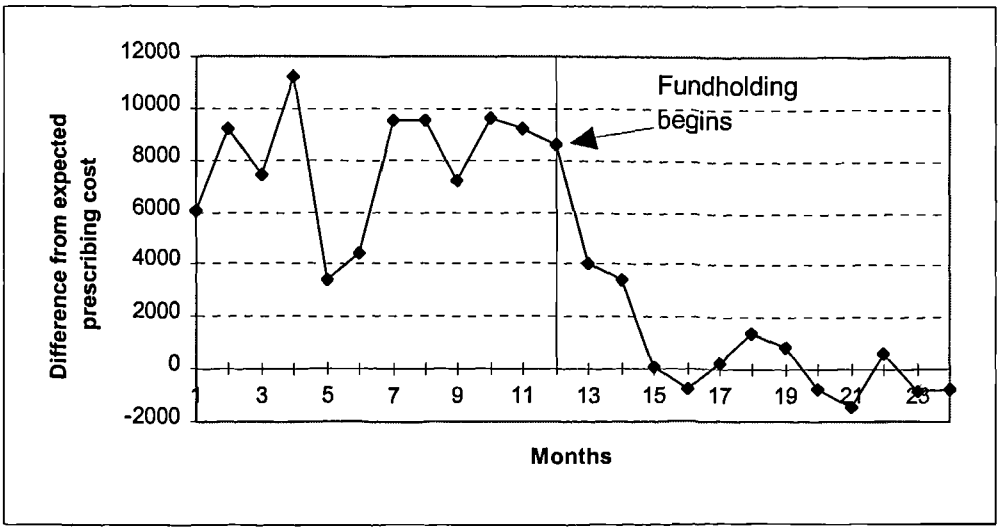



Figure 1.5a Fundholding and difference in prescribing cost (£) of a practice relative to that expected (from district mean cost per / PU).

The scheme should therefore be successful at motivating GPs to change their prescribing cost and rate. My own work bears this out and an example is displayed in figures 1.5a and 1.5b.

* SOS - Survive our Stress. A joint initiative by Wirral Health Authority (Dr C Onion, Mrs Kathy Doran), Wirral Local Medical Committee (Dr Sue Chesters) and Wirral Community Healthcare Trust (Mrs Raie Williams). 1995 - ongoing program.

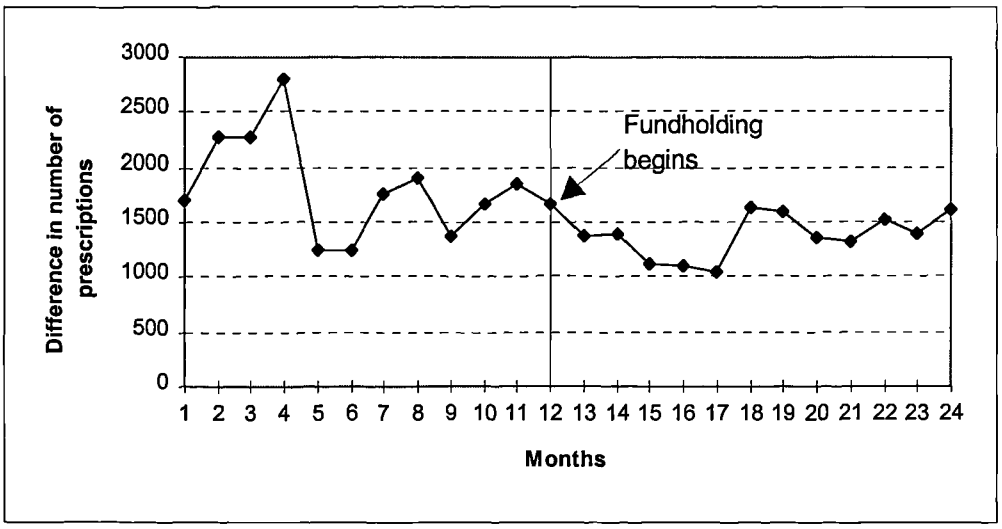



*Figure 1.5b Fundholding and the number of prescriptions issued by the practice in fig. 1.5a relative to that expected (from district mean prescriptions / PU).*

This particular practice in a fundholding consortium was historically 15% more costly in prescribing per head of population than neighbouring practices. The use of expensive antibiotics was particularly high. Prescribing cost fell markedly following fundholding (figure 1.5a). Further analysis showed a remarkable transformation in one particular therapeutic area. The expensive antibiotics had been dropped in favour of penicillin V (the cheapest available antibiotic). The savings had been achieved by using cheaper drugs, not by reducing the prescribing rate (see figure 1.5b). This is consistent with Pitts' and Vincent's finding that it is easier to influence which drugs GPs prescribe rather than their rate of prescribing.

A three year study of the effect of fundholding in the Mersey region by Wilson, Buchan and Walley showed that fundholders' prescribing did change relative to non-fundholders'. The proportion of drugs prescribed generically increased significantly, the rate of increase in prescribing cost was significantly reduced, and there was a small, but significant reduction in prescribing rate (in terms of number, but no data on size) of prescriptions. The effects became less pronounced with each successive wave of fundholder. First wave fundholders were *low volume / high cost per item* prescribers. Second wave were *low volume / mixed high and low cost* prescribers. Third wave fundholders were indistinguishable from non-fundholders.[29] A subsequent study by some of the same researchers on differences in prescribing between the years 1990/1 and 1993/4* showed that there were more powerful influences operating on prescribing than fundholding even though it did seem to be largely responsible for differences (between fundholders and non-fundholders) in the rises in costs between the study years.[30] In a national study of the first 5 years of fundholding it was shown that for each successive wave of fundholders there was a reduction in the growth in prescribing costs (relative to continuing non-fundholders) of about

---

* See appendix A for a critical appraisal of this study.

4%, but this became similar to non-fundholders growth after about 3 years - in other words the effect was modest and was not sustained.[31]

Beyond prescribing, Surender, Bradlow, Coulter *et al* showed that fundholding has not affected the rate of referral of patients to secondary care[32] and a recent BMJ review concluded that the scheme seems to have been ineffective at changing the way doctors practise.[33]

## 1.6 Educational preferences.

From a cognitive point of view, educational interventions can be divided into two categories. *Superficial* techniques, where the information is processed by the recipients in a shallow, sensory way; and *deep* techniques where processing is more effortful, relating the information to previous knowledge and experience to appraise the implications. Examples of superficial methods are lectures and reading, where a passive audience has information transmitted to it. Deep methods are, for example, clinical discussion meetings, audit meetings, small group work, problem based learning,[34] and action learning sets,[35] where a group of doctors and/or others are involved in interactive debate requiring effortful and creative thinking.

A Canadian survey of preferences of doctors regarding different types of continuing medical education (CME) was performed in 1981.[36] The doctors were asked which interventions had most affected their ways of managing patients, they said;

| | |
|---|---|
| reading | 42.5% |
| lecture courses | 18.8% |
| informal discussions | 14.6% |
| formal consultation with specialist | 12.4% |

Superficial methods were clearly considered to be more effective. This led the authors to conclude that an appropriate

development would "...include improving methods of providing physicians with the best information available when it is needed".

A major survey of British continuing medical education by Branthwaite and colleagues in 1988 asked 632 GPs for their opinions upon the usefulness and enjoyment of different types of courses (on five-point scales: 1 = negative, 5 = positive)[37]:

| | | Enjoyment | Usefulness |
|---|---|---|---|
| 1. | Reading | 3.8 | 3.4 |
| 2. | Lectures | 3.5 | 3.4 |
| 3. | Audit or clinical review | 3.3 | 2.9 |

GPs preferred reading best of all, followed by lectures then group work. As in the Canadian study, superficial methods were considered most effective, and were also preferred. The GPs were also asked about which topics they were interested in knowing more about:

| | | *Percentage interested:* |
|---|---|---|
| 1. | Clinical Medicine | 65 |
| 2. | Therapeutics | 54 |
| 3. | Psycho-social problems | 40 |
| 4. | Alternative medicine | 40 |
| 5. | Practice management | 35 |
| 6. | Pharmacology | 28 |
| 7. | Other | 16 |

On the strength of this subjective data it might be tempting to produce a range of literature concentrating on clinical medicine and therapeutics for GPs to choose from. Recently an influential British discussion paper suggested that, "A change in emphasis to self-directed learning based on experience implies a higher level of participation by individuals and the adoption by providers of continuing medical education of appropriate learning formats". The media recommended were reading, reflection and audit (in that order).[38] Superficial methods are preferred by doctors; perhaps, we might conjecture, because less effort is required.

In 1982 a further Canadian study Sibley and Sackett *et al* tested the hypothesis that educational literature on demand was effective at improving standards of care.[39] In a randomised controlled trial doctors who received educational literature on topics in which they were interested in knowing more about showed no subsequent improvement in the quality of their management of subsequent relevant cases. In a second phase the experimental group received educational literature on non-preferred topics. In this situation their quality of care of relevant subsequent cases *did* improve significantly, but there was no spill-over effect into other topic areas. It appears that educational bulletins on relevant, but *not* preferred, topics are effective at

improving clinical practice. The doctors preferred to learn about topics in which they were not likely to improve their management; we might conjecture that they chose topics of which they already had a good grasp, thereby requiring less effort.

In a 1982 American study the effect of a drug bulletin (on the use of oral analgesics in a teaching hospital) showed a significant, *but short-lived*, effect on prescribing. The effect began to wane at five weeks.[40] It would appear that printed material might have a significant impact on prescribing behaviour in the short term, but the effect is not necessarily sustained.

A recent review of international experience in controlling drug expenditure found the available studies suggest that simply providing information on its own does not lead to substantial changes in practice and more active strategies should be considered.[41]

In a Welsh study comparing the effect of feedback by printed material ('prescribing analysis workbooks') and face-to-face discussion visits by a pharmaceutical adviser on antibiotic prescribing it was found that both methods had a positive impact sustained for at least 24 months, but significantly more change was evident after face-to-face visits than after receipt of workbooks.[42]

The only study that I can find that has contrary results indicated that a formal seminar was more effective than small group work was an Indonesian study promoting better management of gastro-enteritis. However, the seminar was (surprisingly) more expensive per doctor than the small group work (normally small group work is more labour-intensive) which begs questions about whether the methods used were comparable to the usual British varieties.[43]

A Dutch review of educational strategies for implementing change in primary care found that individual instruction, feedback and reminders seemed to be the most effective single strategies, and the most effective combined strategies always included individual instruction and feedback with peer review.[44]

In 1995 a systematic review of published evidence on effectiveness of continuing medical education by the University of Toronto discovered 99 trials (covering 160 interventions) that met their criteria for rigorous evaluation (randomised controlled trials and similar trials, repeatable educational interventions aimed at changing clinical behaviour or disease outcomes, studies of practising physicians where objective performance measures were evident).[45] Effective strategies included reminders, patient-mediated interventions, outreach visits, opinion leaders and

multifaceted activities. Ineffective strategies included educational materials, audit, and formal CME conferences or activities.

In 1991 the author studied the effect of a postgraduate lecture on *Helicobacter* eradication therapy for duodenal ulcer.

Study details

**The effect of a postgraduate lecture on relevant prescribing by general practitioners: helicobacter eradication therapy**

Researcher: CWR Onion

- Purpose: to measure (a) the effect on the prescribing of a relevant drug by GPs and (b) the persistence of effect of an hour-long didactic lecture event.
- Subjects: Invitations sent to all 70 general practices in Wirral district in 1991.
- Method: Prospective cohort study - the impact of well conducted lectures by two experts (a Consultant Gastroenterologist and a Senior Lecturer in Clinical Pharmacology) on the prescribing of practices represented at the event were compared with non-attending practices.
- Data: Quarterly (3-monthly) prescribing data (PACT data of all dispensed NHS GP prescriptions of bismuth (tripotassium dicitratobismuthate), a vital component of the recommended,

though novel in 1991, treatment for peptic ulcer treatment, aggregated at practice level.

- Analysis: Ten GPs representing 10 different Wirral practices attended the lecture event (other people were also present). The attending practices issued nearly 3 prescriptions per 10,000 patients for bismuth in the quarter before the intervention, non-attending practices prescribed at a similar rate. Non-attendees' prescribing continued at the same rate, but attendees' practices increased their prescribing of bismuth by 100% in the first 3 months post-intervention (chi squared test P=0.05). However, no difference was discernible in the following quarter or thereafter (figure 1.6a).

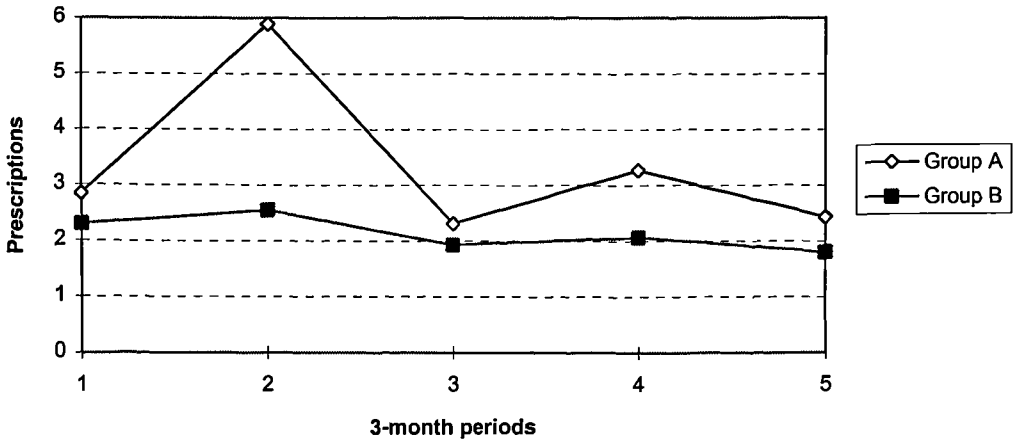



Figure 1.6a Mean tripotassium dicitratobismuthate prescriptions per 10,000 population per practice per quarter in lecture-attending practices (group A) and non-attending practices (group B) following a lecture (period 1) advocating Helicobacter eradication therapy for peptic ulceration.

- The use of practice data rather than individual practitioner diluted the effect. The power of the study (due to low attendance at the lecture event) may have led to type 2 errors in quarter 2, 3, and 4.
- Conclusion: a well conducted lecture on a novel area of treatment can deliver an immediate effect on prescribing behaviour, but the effect can be completely lost after 3 months.

Lectures and literature (i.e. superficial methods) have a certain feature in common, both are effective, but the effect is not persistent. Waning occurs at about 5 weeks and change has reversed by 3 months at the latest. If lectures and literature are

employed in parallel a more persistent synergistic effect might be expected. However, in a Boston hospital in 1982 there was a two-month educational programme involving lectures, bulletins and notices concerning the rational use of gentamicin; this had resulted in significantly improved prescribing of gentamicin which only persisted for six weeks.[46] Lectures and literature cause changes in behaviour that are not persistent even when the two methods are used together.

Medical educationists tend to categorise their interventions as 'active' or 'passive' indicating the level of effort of the recipient. The studies above strongly suggest that active involvement is more effective than passive — i.e. discussions more than lectures, small group work rather than reading, interactive feedback rather than receipt of printed material etc.. But why should this be ?

Educational platforms tend to be descriptive rather than explanatory, for example 'problem based learning' is defined as mastery of knowledge by solving problems so basic information is learned in the same context that it will be used,[47] and 'contract learning' is defined as negotiated independent learning between teacher and learner.[48] Similarly, 'reflective learning',[49] 'clinical audit',[50] 'action learning'[51] and 'learning styles'[52] describe rather than explain educational phenomena. I, however, was interested

to examine the fundamental psychological (especially cognitive) reasons for the success of these methods and to find a testable theory.

Two studies that approached from the learners' perspective described psychological features — superficial versus deep cognitive processing — that were clearly associated with significantly different levels of understanding and memory recall of the same educational material in the same educational environments. In these two studies learners were allowed to select their preferred depth of processing, but if the depth of processing were imposed by the teacher, regardless of the learner's preference, would a similar effect still be observed (indicating a universal feature of human cognition) ?

To examine this I studied the links between this quality of cognition — depth of cognitive processing of messages — and memory recall for medical facts.

## 1.7 Memory.

One of the most important advances in memory has been the depth of processing theory of memory of Craik and Lockhart. This proposes that messages arriving through the senses are processed (analysed) either superficially or deeply.[53] In superficial processing merely the volume, pitch, brightness or similar raw sensory information is considered. Deep processing however, involves analysing the meaning of the message. Associated items, related images, past experiences, implications and semantics are considered. Superficial processing results in memories that are vague and easily forgotten, but deep processing results in memories that are accurate and persistent. The theory has been verified through rigorous experimental testing. For instance, in one study, people who had been asked to make judgements of meaning of a list of words remembered many more of them than other people who were asked to make superficial judgements (e.g. about the size of the letters or number of vowels)[54].

The author has demonstrated a similar effect in continuing medical education where clinicians asked to relate a list of diseases to their previous experiences remembered far more than those given exactly the same amount of time to examine the mere characteristics of the words.[55]

Study details

**Depth of information processing and memory for medical facts**

- Researchers: CWR Onion, PD Slade
- Subjects: 24 medical and pharmaceutical advisers
- Setting: workshops at a national conference in1993
- Method: In the experiment there were two consecutive workshops, twelve participants (medical advisers and pharmaceutical advisers) attended each workshop (a total of 24) without any forewarning of the specific content and procedure. In each workshop the participants took seats at random on either of two sections separated by an aisle. In workshop 1 the participants in one section were designated 'group A'; those in the other, 'group B' . In workshop 2 the labelling was reversed. Participants in group A were intended to engage in superficial processing; those in group B in deep processing. Each group was given a questionnaire consisting of bivalent (yes or no) responses to two simple questions for each of the 30 infectious diseases notifiable by law to the medical officer of environmental health (consultant in communicable disease control). Group A was confined to *superficial*, but precise evaluation. For each notifiable infection participants were asked:

***a) Does it have more than eight letters?*** ***Yes / No***

***b) Does it contain more than three vowels?*** ***Yes / No***

Group B was engaged in *deep* processing. For each notifiable infection participants were asked to relate to their own knowledge and experience:

***a) Does it have a unique diagnostic feature?*** ***Yes / No***

***b) Have you ever come across a case?*** ***Yes / No***

The questions were designed to polarise the depths of processing in the two groups. Under normal circumstances a student is unlikely to avoid all consideration of the meaning of a word, nor is he or she likely to examine the personal relevance of every word. However, if mixed approaches were allowed the results would be inconclusive. A preliminary questionnaire test demonstrated that volunteers took 105 seconds to complete either questionnaire.

- *Part 1*: Participants in groups A and B running concurrently were allowed precisely 105 seconds to respond. Further time would have allowed the participants to rehearse their response and thus confound the experiment. After the completed questionnaires were collected the participants were deliberately distracted by a discourse on a different subject for 60 seconds, to allow any short-term memory to wane (Atkinson and Shiffrin 1968 [56]).

- *Part 2*: Participants in the first workshop were individually given a blank sheet of paper and were asked to write down as many notifiable infections as they could remember. Whilst the full list of notifiable infections was displayed, the participants were invited to total their individual correct answers. Participants in the second workshop assessed their scores in the same manner.
- Results: Group B (deep processers) scored better than group A (superficial processers) in both workshops. None of the participants in the superficial processing group could recall more than half of the 30 notifiable infections. In contrast, 58% of the participants in the deep processing group recalled more than half of the notifiable infections. A Fisher-Irwin exact test on the two by two table (deep / superficial; pass / fail) shows the two-tailed probability of there being no difference in recall between the two levels of processing as 0.005. The odds ratio is infinite with 95% confidence limits of 2.226 and infinity. It is therefore significantly unlikely that the relatively high rate of memory recall in group B had occurred by chance.
- Conclusion: deeply processing medical facts significantly improves memory recall.

The levels of processing theory also indicates how recall may be improved by rehearsal of the message, or repeated processing. There are two types of rehearsal. *Maintenance rehearsal* merely repeats the same processing and is not associated with better recall, but *elaborative rehearsal* involves increasingly deeper processing and is associated with better recall. Bulletins and lectures, being a one-way transmission of messages to a passive audience, are not likely to invite deep processing of their messages. The depth of processing theory predicts that memories formed will be short-lived. The studies on the effects of lectures and bulletins above support this hypothesis. If however, a bulletin or lecture invokes a slight degree of deep processing then repeating the message through a series of bulletins and lectures will result in elaborative rehearsal and improved recall. The more the message is repeated, the stronger and more persistent the memory. Indeed, Soumerai and Avorn in a Harvard study of medical education in 1987 found that brevity, repetition and reinforcement are important components of effective education.[57] Practitioners are more likely to remember repeated messages, but that does not necessarily mean that they are more likely to comply with them.

## 1.8 Deep processing and attitudes: the Elaboration Likelihood Model of persuasion

There has been a common finding in experiments studying the effect of repeating messages to change people's behaviour. Up to a point people increasingly agree with a message, but beyond that point they increasingly disagree with it. An explanation is suggested by a theory that links communication, attitudes and behaviour, this is the *elaboration likelihood model (ELM)* of persuasion.[58] This theory proposes that deep processing* of the meaning of a message content will result in an attitude towards the issue of increased strength. If the thoughts generated are in agreement a positive attitude is formed, if the thoughts are in disagreement with the message the attitude will be negative. If the thoughts are neither favourable nor unfavourable the individual retains his original attitude. Repeating a strong argument can be expected to increase agreement and create a more positive attitude. This is what the evidence shows up to a point, but why does agreement fall away after too much repetition? The answer lies in a phenomenon called *reactance* where an individual perceives that a message is being forced on him and deliberately adopts the contrary stance.[59] The experiment that clarified this two-stage effect was performed in 1979 by Petty

and Caccioppo.[60] A series of reasonable arguments recorded on audio-tape were played repeatedly to a group of subjects who were asked how much they agreed with them. There was increasing agreement after each repetition up to *three* repeats. Thereafter agreement dropped away progressively. Deep processing of a reasonable message results in better recall and a positive attitude. This effect is proportional to the degree of deep processing. However, if repetition is used to facilitate deep processing then there will be rejection of the message after three close repeats.

---

*'Deep processing' and 'elaboration' are synonymous.

## 1.9 Deep processing and behaviour: two routes of persuasion.

The elaboration likelihood theory of persuasion postulates that superficial processing involves a different *quality* of thinking to deep processing. So marked is the difference in quality that the model proposes that there are two distinctly different routes to persuasion, the central route which employs deep processing, and the peripheral route which employs superficial.[61] Their characteristics are:

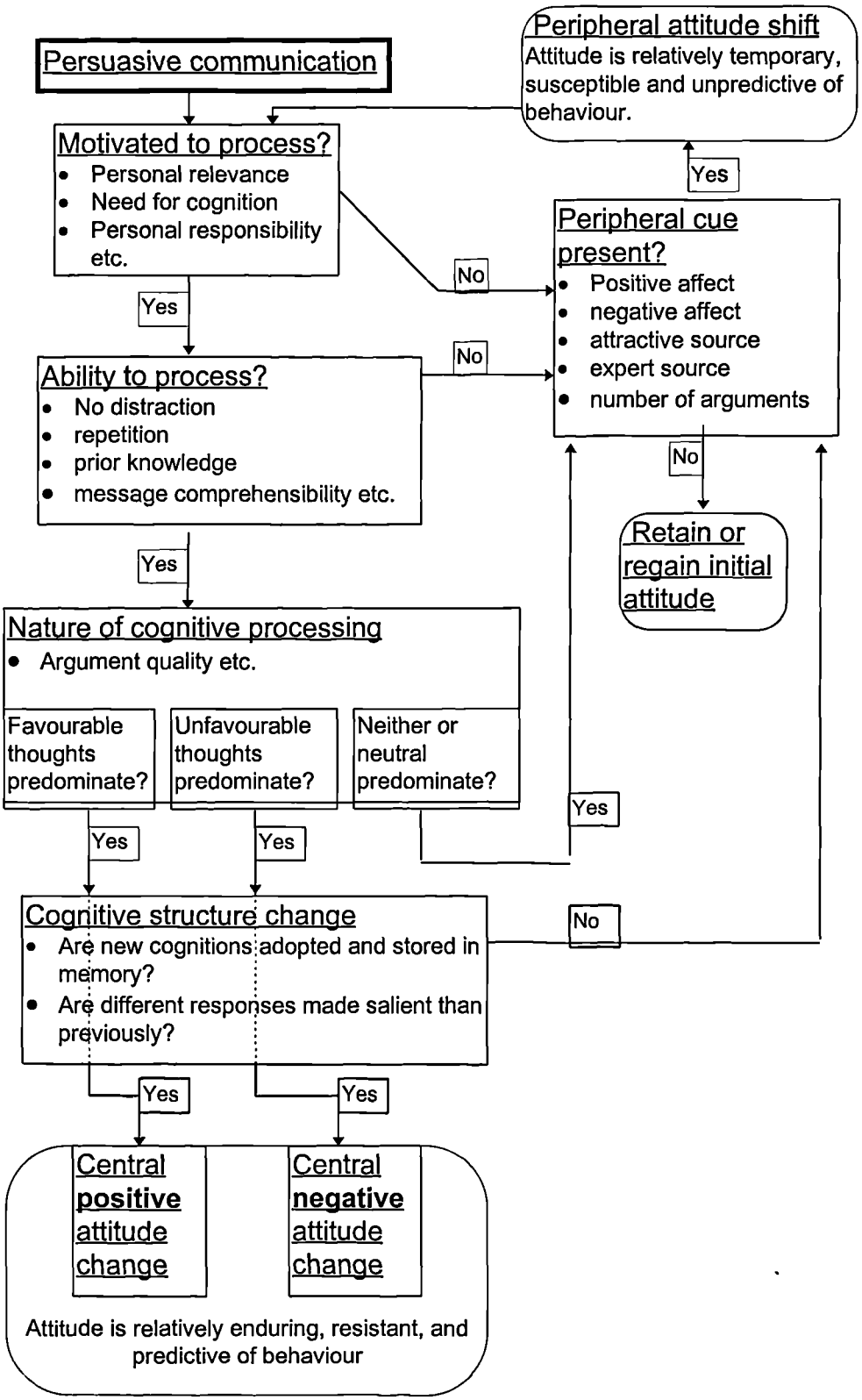
*Central Route:* there is effortful scrutiny (deep processing) of the true merits of the information presented. In these circumstances *the attitude gained is persistent, resistant and highly predictive of behaviour.*

*Peripheral Route:* there is no effortful scrutiny and the resulting attitude gained is due to some simple reflex sensory cue (superficial processing), such as the attractiveness (an affective cue) or expert status of the source (a simple implication). In these circumstances *the attitude gained is temporary, susceptible and unpredictive of behaviour.*

The peripheral route (for superficial processing) is quick and easy, it enables people to act upon an enormous amount of daily information efficiently. There are dangers in total reliance on this

route and the mind appears to have a safety mechanism; attitudes formed by the peripheral route are *temporary, susceptible and unpredictive of behaviour.* The evaluations formed are short-lived, easily changed, and not automatically followed by consistent action. The process is displayed overleaf.

Figure 1.9a (overleaf) — Central and peripheral routes to persuasion; adapted from Petty and Caccioppo (1986). Information shown in bold box, attitudes in rounded boxes.


Persuasive communication
Motivated to process?
• Personal relevance
• Need for cognition
• Personal responsibility etc.
Yes
No
Ability to process?
• No distraction
• repetition
• prior knowledge
• message comprehensibility etc.
Yes
No
Nature of cognitive processing
• Argument quality etc.
Favourable thoughts predominate?
Unfavourable thoughts predominate?
Neither or neutral predominate?
Yes
Yes
Yes
Cognitive structure change
• Are new cognitions adopted and stored in memory?
• Are different responses made salient than previously?
No
Yes
Yes
Central positive attitude change
Central negative attitude change
Attitude is relatively enduring, resistant, and predictive of behaviour
Peripheral attitude shift
Attitude is relatively temporary, susceptible and unpredictive of behaviour.
Yes
Peripheral cue present?
• Positive affect
• negative affect
• attractive source
• expert source
• number of arguments
No
Retain or regain initial attitude

People receive thousands of verbal and written messages every day. It would be impossible to analyse each one in depth because of the sheer overwhelming volume of data so most are automatically processed via the peripheral route. Instead of examining the deep meaning, considering the implications of following the course of action suggested by the message, and relating it to previous knowledge and experiences people understandably take a short-cut. People look for simple associations, features which have always, or nearly always, been related to success or failure in the past, and act accordingly.[62] Examples would be, the unquestioning belief of messages issuing from a superior, or expert, or media personality. For example, "He's the boss, he's in a position to know", or "She's the expert in this field, if she thinks that every mild hypertensive should receive this new drug then it must be right", or "If Sir Arthur Conan Doyle is convinced by the photographic evidence that fairies exist, that's good enough for me". Sometimes there is even less rationality in the link, the association being purely affective or symbolic. For instance, "This very *pleasant* salesman says that this is the best deal he can give me on the car so it must be" or "This petrol company always uses *rural scenes* in advertisements so they must be environmentally friendly and can therefore be trusted". It is tempting to assume that this route is without any merit. It is in

fact the basis of empirical science. However, it is not meritorious to entirely rely upon repeated associations that may merely be the result of chance (see chapter 1 leading quotation; Hume). There is also merit in considering the underlying reasons for phenomena (see chapter 2 leading quotation; Leibniz).

Deep processing requires a good deal more mental effort and concentration and is therefore often avoided. Of the myriad messages received each day, all being superficially processed, some are selected for special attention. These are messages that are likely to have a bearing on personal goals or responsibilities. In other words they are personally *relevant*, *interesting*, *related to important duties*, or *likely to be the subject of an expected enquiry*. If the conditions are right (e.g. no distractions), the message comprehensible, and the required prior knowledge possessed, deep processing will take place. The message is evaluated against experience, beliefs, theory, soundness, related topics, categorisation, and its implications considered. Deep processing is relatively cold and objective. At the end of the process an evaluation takes place. If the thoughts are favourable or unfavourable, a (positive or negative) attitude results and is memorised. The persistence of the memory is proportional to the depth of processing. The degree of polarisation of the attitude is

also proportional to the depth of elaborative processing. Therefore, the central route results in an attitude that is relatively *persistent, resistant to counter-persuasion and is highly predictive of actual behaviour.* The process is summarised in figure 1.9a.

The elaboration likelihood model is supported by experimental results that verify each individual component,[63] and it features in current social psychology textbooks,[64] but has not been extensively tested in the field. There is a recognition that its application to clinical psychology must be scientifically explored.[65] The application of the ELM theory to the field of medical education has not been recorded or tested, though it is appropriate that it should be because the substantial effect of learner-preferred deep processing itself on subsequent levels of understanding and memory recall has been demonstrated (as described in sections 1.6 and 1.7 above). Deep processing is consistent with active educational methods and superficial processing with passive methods, though I will adopt the cognitive terms in this thesis because they represent the aims of different methods of education, rather than using active and passive which are merely descriptive terms.

Testing the elaboration likelihood model in the environment of continuing medical education is one of the objectives of this thesis.

---

# Chapter 2

# Feedback, detailing and formularies

> *The idea of cause and effect is deriv'd from experience, which informs us, that such particular objects, in all past instances, have been constantly conjoin'd with each other: and as an object similar to one of these is suppos'd to be immediately present in its impression, we thence presume on the existence of one similar to its usual attendant.*[66]
>
> David Hume (1739)

## Summary of chapter

Feedback that has effected persistent favourable change in behaviour has motivated practitioners to process issues deeply. Drug formularies (prescribing guidelines) are only effective if the recipient practitioners are involved in deep processing of their content. Another effective way of achieving deep processing is through engaging practitioners in discussion on clinical topics in face-to-face interview; this too has persistent effect. The phenomenon is exploited by commercial promoters of pharmaceutical products and their sales personnel who visit practitioners, possibly compromising the efforts of medical educators who promote more conservative approaches to prescribing; new drugs are most heavily promoted before their

true position in treatment and clinical experience has been established. However, deep processing evaluates messages rigorously and tends to identify weak arguments.

## 2.1 Feedback and reward.

The classical theories of Skinner and other behavioural psychologists on the relationship between feedback and behaviour change (operant conditioning), predicts that feedback on clinical performance will produce a beneficial change in physicians' behaviour.[67] The results however, are equivocal. In New Mexico in 1971 a group of physicians approved a set of criteria for the injecting of drugs.[68] The performances of physicians were audited and any injection events not complying with the criteria were denied payment. The injection rate fell by 60% in two years. This model established punitive audit in the USA. However, evidence of subsequent successes for audit in America is hard to find.[69]

Since 1988 quarterly feedback has been sent to every English and Welsh GP on personal prescribing rate and cost (Prescribing Activity and CosT [PACT]) by the Prescribing Pricing Authority. These allow cross-sectional comparisons with the district average and longitudinal comparison of personal performance. They appear to have resulted in no significant reduction in the rate or cost of prescribing.[70] PACT was given a sting In 1991; it was supplemented with monthly feedback that compared personal performance with target expenditure. For most

practitioners the target was a fictional budget, the Indicative Prescribing Amount (IPA). Practices who varied significantly from this amount would have to justify the variance to the Family Health Services Authority (FHSA) medical adviser.

According to my analysis of national data the intervention in itself had no significant effect (figure 2.1a).

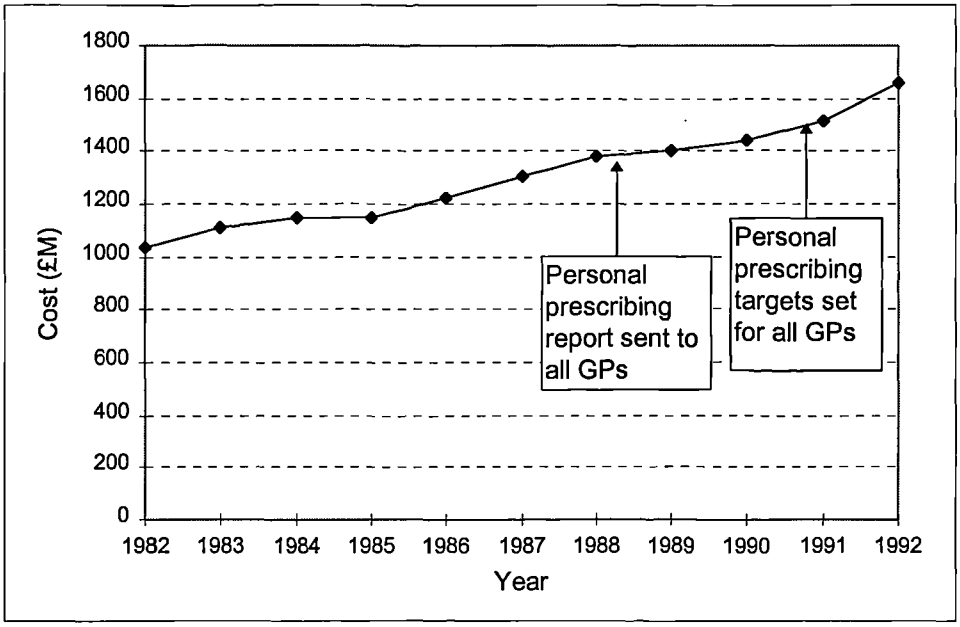



Figure 2.1a NHS prescribing cost (corrected for inflation at 1982 prices): English general practice 1982 to 1992. [71]

However, I have found that individual meetings between medical or pharmaceutical advisers and GPs have produced behaviour changes regarding issues discussed (figures 2.1b and 2.1c). The examples in the next two figures below have been selected because they varied from the district mean by

considerable degrees for the variable under test prior to intervention. The remedial changes are therefore easily discerned. However, improvements can even be achieved in average and better than average practices.

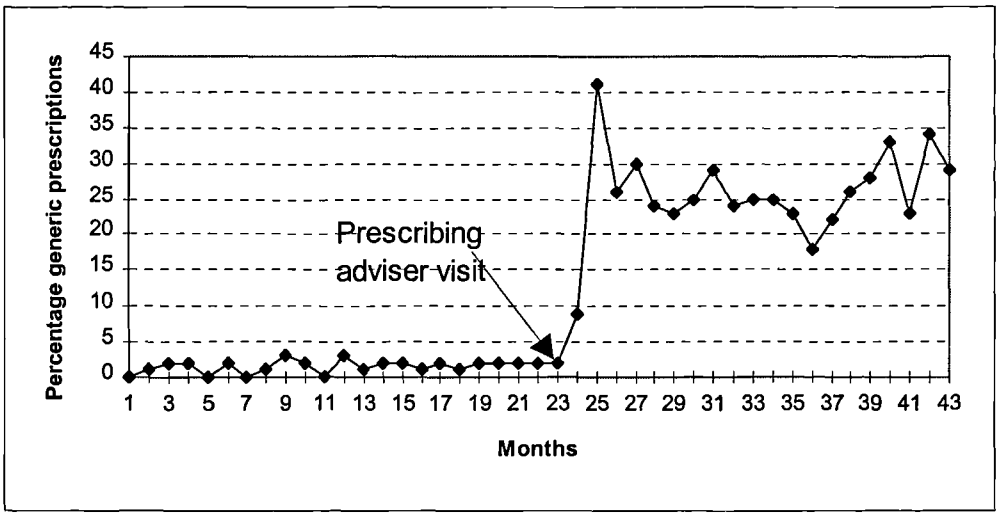



Figure 2.1b The effect of the visit of a pharmaceutical adviser to a general practitioner specifically discussing generic prescribing. Note the persistence of the effect.

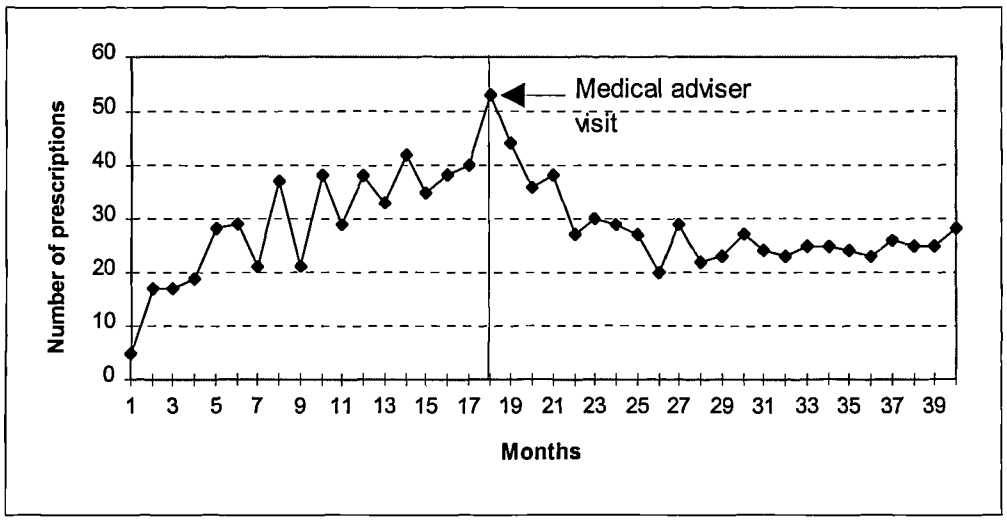



Figure 2.1c The effect of an interview between medical adviser and single-handed general practitioner with discussion of appropriate use of lipid lowering drugs. Note the persistence of effect.

In a national scheme, other GPs (Fundholders) were given a real financial budget with which to cover the costs of their prescribing and were allowed to invest any underspend in their practices. Fundholders also enjoyed other benefits such as greater influence over hospital services.

However, in a local pilot scheme the author and a colleague studied the effect of a combination of motivational factors.

Study details

**Combining financial and intrinsic reward: a pilot prescribing incentive scheme for non-fundholding general practitioners**

Researchers: CWR Onion, Clare E Dutton.

- Subjects: All non-fundholding practices in Wirral district were invited to participate in a prescribing incentive scheme. The first 10 were selected. A further 10 non-fundholding practices were selected as a control group to match the features of the 10 with regard to size of practice, number of partners, geographical locality and overall prescribing performance (cost per prescription and rate {number of items} per prescribing unit {over 65s counted thrice}).
- Design: Open prospective matched controlled trial.
- Method: We presented 10 interested practices with a financial incentive (£3,000) to achieve a challenging, but achievable agreed financial target on their prescribing cost (negotiated with the practice). They were also required to demonstrate, through clinical audit in the same period, a successful improving prescribing initiative in the practice.

—The intention was to combine both extrinsic (financial reward) and intrinsic motivation (better medicine) so that any financial savings were not achieved by reducing the quality of prescribing. No feedback on prescribing was given to the incentive scheme practices during the study period, only upon completion. After negotiating the target, the only interaction with the researchers was a compulsory audit meeting arranged by the MAAG where practices exchanged experiences on the various change strategies they had employed to attempt to achieve their targets.

- Data: PACT (Prescribing Pricing Authority database of all NHS GP prescriptions dispensed) for the 3 months prior to invitation of practices (Sept-Nov 1992 = baseline) and the 3 months post-intervention (March-May 1993). Subjected to paired analysis.
- Results: All 10 incentive group practices completed audit cycles and 6 achieved their financial targets. One practice refused the financial incentive from the outset, being content with trying to achieve a successful audit. Favourable changes were observed in the incentive practices, but not in the matched control practices.

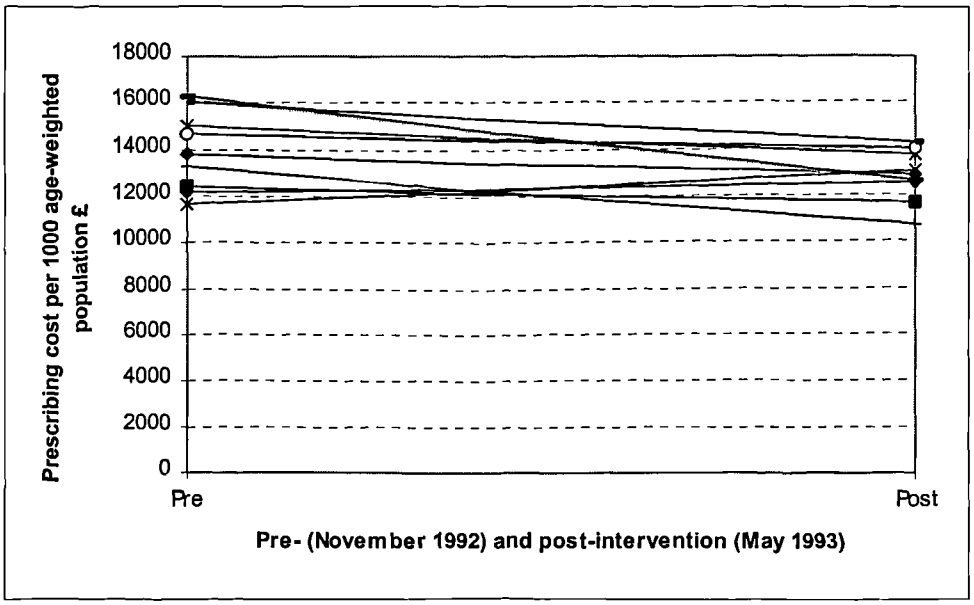



Figure 2.1d Changes in standardised prescribing cost in the pilot incentive scheme practices. Changes are generally downward.

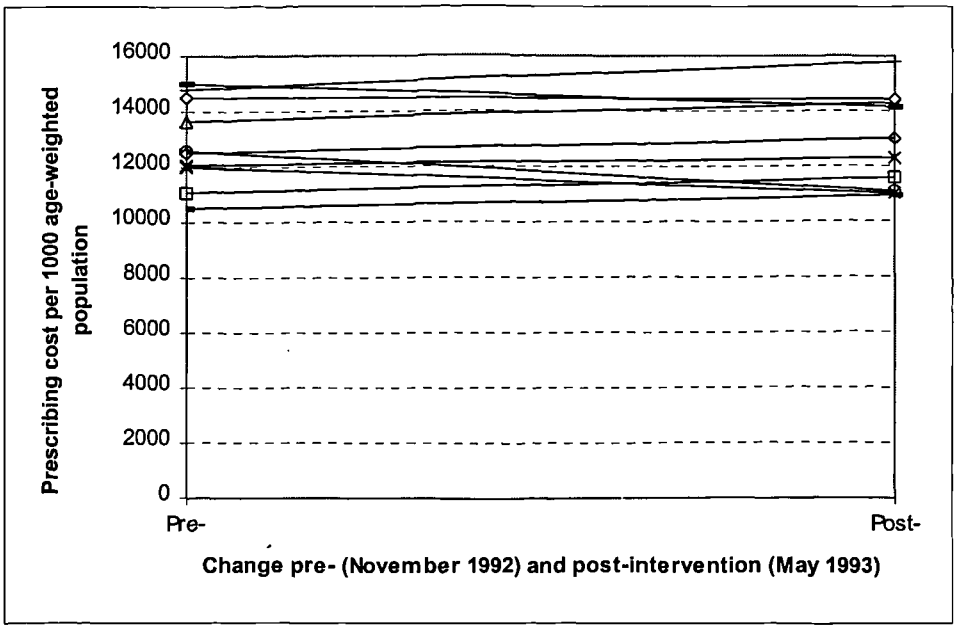



Figure 2.1e Changes in standardised prescribing cost in the incentive scheme matched control (no intervention) practices. Changes are generally upward.

—

—The 10 practices in the incentive scheme had significantly reduced prescribing cost compared with controls (Mann-Whitney U test, exact 2-tailed $P < 0.05$, median difference in cost per 1000 PUs = **-£1,275**, 95% CL = £60 to -£2,310). The proportion of drugs prescribed generically was increased ($P = 0.008$, median difference = **+4.2%**, 95% CL = 1.1% to 8.6%). However, as with the fundholding scheme (see chapter 1, table 1.5b), the *rate* of prescribing was not significantly affected ($P = 0.39$, median difference = **-56** prescriptions per 1000 age-weighted population, 95% CL = 32 to -193). Clinical audit showed most changes were in repeat prescribing, particularly generic substitution, and could therefore be expected to continue beyond the study period.

- Conclusion: A prescribing incentive scheme combining financial and intrinsic reward can significantly reduce the cost of prescribing and significantly increase the proportion of drugs prescribed generically over a period of 3 months. There was no evidence, in our study that prescribing quality was compromised in order to achieve the savings.

Incentive schemes of this sort have now been adopted nationally. Non-fundholders are offered a purely financial incentive to reduce their prescribing costs. By achieving an agreed lower target cost figure they receive a substantial payment (£3000). The downward change in prescribing expenditure amongst the NHS prescribing (purely financial) incentive scheme general practitioners is said to have been dramatic,[72] but has not been rigorously evaluated.* The cynical effect I observed in one Wirral practice that embarked on the NHS scheme suggests the financial impact could be short-lived (figure 2.1f).

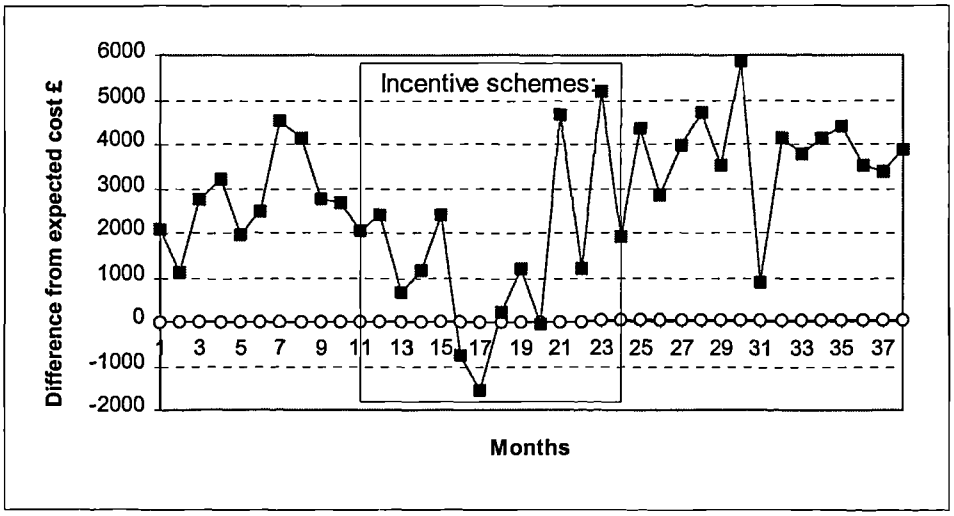



Figure 2.1f The effect of the NHS prescribing incentive scheme on a high cost two-handed Wirral practice. The performance-measured period is marked by a box. Note the marked drop in expenditure relative to that expected from other practitioners in the district which reverses as the incentive payment target is achieved.

* See appendix A for a critical appraisal of the paper referred to here.

The national scheme provides only extrinsic motivation and could therefore be expected to show both a short-term effect and waning enthusiasm for subsequent years' schemes.

Direct, distinctive, and real financial reward or penalty tied to performance has had a significant impact upon behaviour. However, abstract feedback based upon notions tied to information (e.g. personal prescribing data reports and indicative prescribing cost targets) has had no impact. Fundholders can easily identify cause and reward; for example switching peptic ulcer patients from ranitidine to cimetidine has resulted in a predictable financial reward. For other GPs the feedback has been intangible, difficult to tie to a specific event and therefore failed to change practice.

Prescribing performance and other aspects of clinical management are linked. A study at Whittington Hospital has shown that GPs whose prescribing complied well with the local antibiotic policy were significantly different in terms of proportion of patients under 5 years old, the (larger number) of urine tests requested - particularly in those under 5, and in the number of high vaginal and cervical swabs taken.[73] It is therefore unsurprising that the effects observed on prescribing can be

observed in other areas of clinical management in general practice.

In a Dutch study published in 1992 biannual feedback from the laboratory to doctors regarding the appropriateness of the investigation in the light of the clinical details on the form reduced the volume of investigations discussed for several years in comparison to a control laboratory.[74] In an extension of this approach the researchers performed a randomised balanced incomplete block study on GPs all using the same laboratory to examine had any effect on the appropriateness of tests requested.[75] Routine feedback reduced the quantity and improved the appropriateness of laboratory test requests. Another Dutch study of feedback on cervical smear specimen quality examined the effect of remoteness of feedback from the related event.[76] Feedback produced improvements amongst gynaecologists and other doctors performing a large number of smears, but there was no improvement amongst GPs. The authors concluded that the difference must be due to the smaller amount of feedback and the interval between receiving the feedback and being able to act on it. It might be many days between feedback and the next smear. To be effective feedback must be distinctive, attributable to a specific intervention, close in time to the event and to the next

opportunity to practise the relevant change. Also, feedback will only bring about change if it affects the achievement of a personally relevant goal (see chapter 1.5). A subsequent study of individualised feedback on diagnostic tests involved recommendations attributable to specific cases or diagnostic topics. Improved performance amongst the GPs resulted. No financial reward or penalty was involved whatsoever.[77] Effective feedback represents persuasion by the central route, the recipient is motivated by the personal relevance of the information to deeply process the message. Changes in attitude formed are therefore persistent, resistant and followed by appropriate behaviour changes (see chapter 1.9).

## 2.2 Face to face interviews: detailing 1

Written information (such as a drug bulletin) invites superficial processing and tends to result in short-lived changes in behaviour. Typically, the bulletin looks authoritative, exudes expert opinion, and contains a plethora of heavily referenced arguments. The characteristics are peripheral cues and lead to changes in attitude that are temporary, susceptible, and not necessarily accompanied by consistent behaviour. What is the effect of using literature to excite interest via the peripheral route followed quickly by a central route approach? For instance, what would be the effect of a face to face discussion on the topic with a pharmacist? The pioneering work in this area was performed by an American physician and pharmacist research partnership in 1983.[78 79] Doctors from four States were divided into three classes. One class acted as a control, a second received printed material only (either a bulletin or a sophisticated glossy 'unadvertisement'), the third received printed material followed by educational visits from pharmacists. The literature and visits promoted alternative therapies to cephalexin, dextropropoxyphene, and peripheral vasodilators. "Physicians who were offered personal visits by the clinical pharmacists together with a series of mailed 'unadvertisements' reduced their prescribing of the target drugs by 14% as compared with controls ($p = 0.0001$)... No such change

was seen in physicians who received mailed printed materials only". The intervention lasted 4 months and the effect at least a further 5 months. There was no difference between the sub-classes receiving either glossy 'un-advertisements' or drug bulletin style material. Doctors were offered an 18 minute face to face interview with a pharmacist covering the three target areas. The topics and specific problem cases were discussed. Superficial processing had been used to excite interest in the topics and was followed by persuasion through deep processing; the physician was actively engaged in discussion and was asked his opinion at several points. A randomised controlled trial in Leeds in 1992, used a similar technique. A pharmacist promoted alternative non-steroidal anti-inflammatory drugs by means of a 'sales' interview. The pharmaceutical company sales representative model was adopted; the potential 'buyer' was encouraged to deeply process the merits of the products in question. A significant change in the prescribing of ibuprofen which persisted for at least five months resulted ($p < 0.005$).[80] The American study adopted an educational deep processing model, the British a sales deep processing model yet the results were the same. Perhaps it is not the model that counts, but the route of persuasion. Furthermore, medical education is following the famous case-study approach of Harvard Business School[81], by developing problem based learning

in medical schools. Problem based learning and other deep approaches are known to be more effective for medical students than traditional superficial methods.[82] Businessmen have discovered that 'involving' their staff in relevant decision-making improves output.[83] In communications *interaction* is superseding pure *transmission,* for example phone-in radio and television programmes.

The singlemost important and unifying principle is that the successful persuader / educator / manager influences his audience through deep processing.

## 2.3 Drug formularies.

In Runcorn, Cheshire in 1981 Green, a research pharmacist, helped a large general practice develop its own formulary.[84] An informal group was formed consisting of the GPs with clinical pharmacologists, pharmacists and members of the primary care team. The group reviewed clinical management issues, exchanged ideas, and endeavoured to develop a more logical approach to prescribing. The result was a significant change in prescribing behaviour towards recommended drugs in at least five therapeutic areas (P ranged from < 0.01 to < 0.001). The changes persisted for at least six months.

In 1987 in Montrose, Scotland the doctors in a general practice together with a clinical pharmacologist developed their own formulary and reduced prescribing costs by 10%.[85] In 1985 Van Zwannenberg studied the educational effect of a prescribing formulary discussion.[86] twelve young Newcastle upon Tyne GPs discussed a local formulary, but were not issued with a formulary, nor did they participate in its development. Their compliance with the formulary was measured before and six months after the educational meeting. The number of initial drugs (i.e. not repeat prescriptions) prescribed that were in the formulary rose from 73% to 83% (P < 0.001). This sort of evidence led McGavoch to

comment in the British Journal of General Practice, "...practice formularies, though difficult to implement, are secure stepping stones to the firm ground of rational prescribing".[87] Within a year his own research investigations cast some doubt on this.[88] A questionnaire survey of 132 practices in Northern Ireland identified those which possessed a formulary. There was no correlation between possession of a formulary and range of drugs prescribed or prescribing cost. The beneficial effects of a formulary reported in the earlier studies were not due to the possession of a formulary. The Newcastle group were not given a formulary and yet their prescribing improved, and the N. Ireland GPs had formularies, but did not use them. Involvement in formulary development is not necessary to guarantee compliance either. The Newcastle GPs changed their prescribing appropriately and yet were not involved in development. *The important factor is a interactive discussion about the issues*; which the 1981 Runcorn, the 1987 Montrose and the Newcastle studies above reported that GPs had been engaged in. This association is confirmed by details which are almost lost in the body of the Newcastle text: After a general discussion about formularies there followed "...a more detailed debate about the advantages and disadvantages of generic prescribing" and a follow up study revealed..."that most of the effects of the intervention had disappeared within two years,

though an increase in generic prescribing persisted". The Newcastle GPs were persuaded to prescribe generically through deep processing of the arguments. As the elaboration likelihood model predicts, their subsequent attitudes towards generic prescribing proved persistent, resistant (in an environment of pharmaceutical company counter-persuasion) and were highly predictive of behaviour.

## 2.4 Pharmaceutical company strategies

GPs' preferred form of education is reading. According to Branthwaite *et al* in 1988, reading rates highest in terms of both usefulness and enjoyment. Their study also asked which periodicals the GPs read regularly.[89]

The author examined whether pharmaceutical companies' advertising was related to the popularity of publications.

Study details

**Relation between proportion of promotional content and readership in journals read by general practitioners**

Researcher: CWR Onion

- Purpose: To assess the accuracy of pharmaceutical company advertising in medical journals read by GPs.
- Subjects: the 8 most popular journals reported by GPs in an RCGP study of continuing medical education by Branthwaite *et al* (see above).
- Design: Observational survey.
- Method: In the last week in April in 1993 I took the current issue of the publications most read by GPs, calculated the percentage of total page area that was occupied by commercial drug advertisements and plotted readership against page area (figure 2.4a).

- Results: Advert area matched readership to a significant degree (two-tailed p = 0.001, Pearson's product moment correlation coefficient = 0.917).

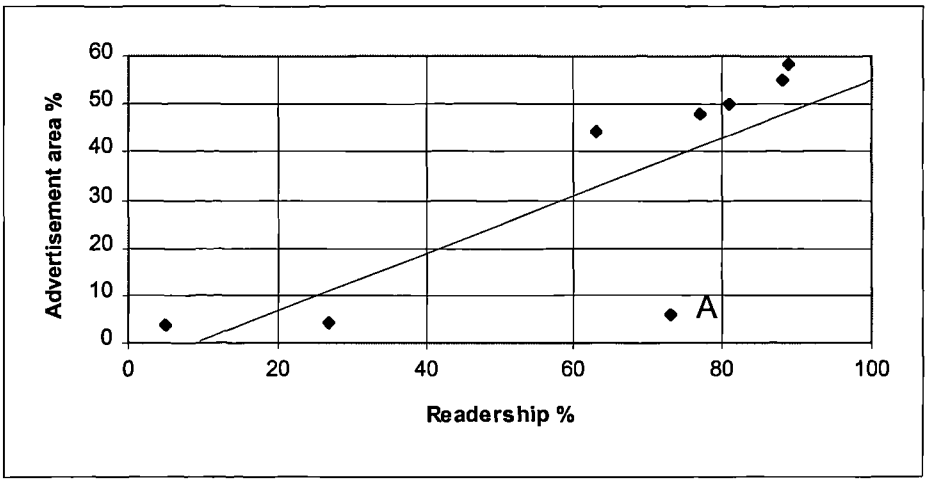



Figure 2.4a Area of journal devoted to pharmaceutical company advertisements plotted against readership in April 1993 for medical journals / periodicals read by British general practitioners. The regression line is displayed. A is the BMJ which prints most advertisements in its supplements.

—In the more popular periodicals over half of the total page area was taken up by commercial drug promotional information.

- Conclusion: Pharmaceutical company marketing is widely believed to be amongst the most sophisticated in the commercial world. The accuracy of targeting of drug promotional messages to the relevant audience tends to confirm this.

In many cases the messages from pharmaceutical companies will counter the medical profession's drive to encourage rational prescribing. The drugs promoted are new, relatively untested, expensive compounds. Pharmaceutical companies do not promote products that are off-patent. Development costs have to be recouped quickly before expiry of patents. New drugs do not necessarily represent an advance in treatment. There are many examples of licensed drugs being withdrawn or restricted because of side-effects only apparent as use became more widespread. For example, in 1988 xamoterol was licensed for all grades of heart failure, and it was quickly and heavily marketed to GPs. Evidence of its toxicity accumulated rapidly and In 1990 the Committee on the Safety of Medicines released a recommendation that treatment with xamoterol be initiated only in hospital and be strictly limited to chronic mild heart failure.[90] This would be of little concern If drug company influences were restricted to printed matter because literature alone is ineffective in the long term. However, "Marketing by pharmaceutical companies relies heavily on the use of printed advertisements *together with face to face visits by representatives"* (my italics).[91] This offers opportunities for pharmaceutical sales representatives to engage prescribers in deep processing of messages; the Elaboration Likelihood Model

of persuasion psychology predicts that following deep processing persistent, resistant changes in attitude and consistent behaviour are likely.[92]

## 2.5 Pharmaceutical company representatives' visits: detailing 2

It was estimated in 1986 that the pharmaceutical industry spent over £5000 on sales promotion per GP per year.[93] The industry must think this a good investment, but GPs do not; only 34% feel that drug sales representatives are a source of knowledge of drugs.[94] However, in a study of the introduction of temazepam, contact with the pharmaceutical company representative was the most reliable predictor of favourable reception of the new drug.[95] The GPs clearly underestimated the influence that pharmaceutical company marketing had on them.

The main pressure from the drug industry is for GPs to adopt new drugs. 5.4% of initial (non-repeat) prescriptions, principally antibiotics and analgesics, are new to GPs.[96] The GPs thought that they were influenced by drug sales representatives where perceived therapeutic risk or disease importance was low. The Elaboration Likelihood Model of persuasion predicts that where relevance and personal responsibility are low, the recipients will not be motivated to deeply process information and so will develop unreliable attitudes.

The author analysed the antibiotic prescribing data from a locality cluster of 18 general practices over a three month period

in winter 1992 (described in more detail in chapter 1.2). The population covered was 58,000. Prescribing of new antibiotics is much more variable than established antibiotics. The coefficient of variation is an index demonstrating the spread of data relative to the mean. A figure over 100% indicates that the standard deviation exceeds the mean.

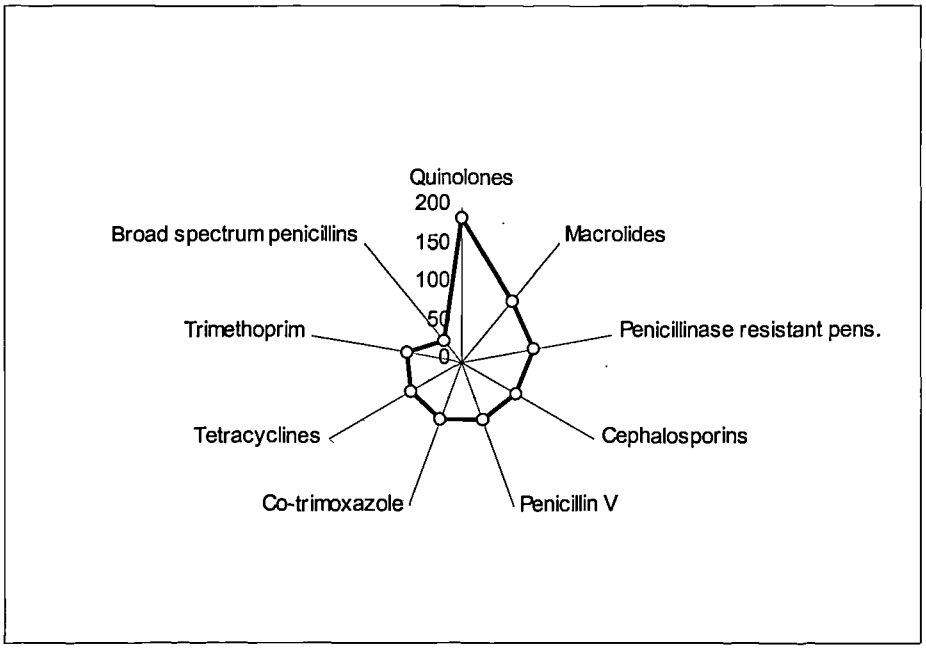



Figure 2.5a Coefficients of variation (%) in prescribing of different antibiotics across 18 general practices. There is more variation in the newer classes of antibiotics.

There is at the extremes a coefficient of variation of 180% for quinolones and 35% for broad spectrum penicillins. The effect may be due in part to the effect of small numbers in the newer classes, but this not the complete explanation. For example cephalosporins are prescribed about as much as tetracyclines (34

and 28 prescriptions per 1000 population), but are subject to 11% more variation.

Although the new drugs only account for 5.4% of initial prescriptions their cost is disproportionately high. GPs have little motivation to deeply process the evidence presented for many new drugs because of the minor nature of the target diseases, but this has significant adverse economic consequences because minor illnesses are the commonest.

Another psychological method employed is the achievement of *commitment,*[97] where the subject is encouraged to state that he will use the product. A public statement of intent also makes an individual more resistant to counter-persuasion, hence the pharmaceutical company representative's badgering at the end of an interview, "You will try this drug on one or two of your patients, won't you doctor?".

## 2.6 Validation of new drugs

New drugs are granted a licence by the Medicines Control Agency after being tested on perhaps less than 1,500 patients, and demonstrating effectiveness relative to a placebo and safety in use. They are introduced without any evidence of effectiveness relative to competing products and little or no adverse effect notification information as this takes time to accumulate. Many of the compounds are marketed directly to GPs who are likely to have too few personal experiences of the drug to notice infrequent adverse effects. The drug salesmen will push all the theoretical advantages their drugs have over the opposition. In the absence of clinical evidence to the contrary how can GPs disagree? Also, the fact that most of the older rival drugs will have accumulated a long list of adverse reactions through the Adverse Drug Reaction reporting scheme can be exploited to make the GP feel uncomfortable about continuing to prescribe older rival drugs. Various well-developed psychological techniques are employed, such as 'yes-chains' where a series of questions are asked which are contrived to force the GP to voice a commitment to prescribing the drug.[98] Verbal expression of commitment produces a compulsion to prescribe[99]. Drug names are carefully chosen to induce a favourable attitude via superficial processing, and will have less than six syllables[100] in order to ensure memorisation.[101]

The original research and development work will only have been allowed to progress because the medical and marketing directors of the pharmaceutical company were convinced that there would be a large market for the product.[102]

Doctors might think that any pharmaco-economic evidence presented will be objective. Not so it seems, for this is likely to be handled by the marketing department.[103] Company post-marketing surveillance studies might be expected to give sure and early warning of any unforeseen problems, but they "... have made only a limited contribution to the assessment of drug safety, principally because of weak study designs and difficulties in recruitment".[104]

Pharmaceutical company marketers have also infiltrated the continuing medical education system[105] and consensus conferences.[106] In the presence of so much sophisticated, biased information one might expect all initially prescribed drugs to be new drugs, but only 5.4% are new.[107]

To be persuasive by the central route a message must be strong, the stronger it is the more persuasive it is. The Elaboration Likelihood Model defines a strong message as one that when thought about the conclusions drawn are favourable. Conversely, weak messages give rise to unfavourable thoughts when scrutinised. In other words, strong messages tend to be logical;

either the conclusion follows from the premisses, or the information required to assert the conclusion is present.[108] Furthermore, if the message is strong, increasing the number of arguments contained in it enhances its persuasive power. If it is weak then increasing the number of arguments reduces its power. With weak arguments pleasantness, attractiveness and authority are influential, and are even more effective when personal relevance of the topic is low. Pharmaceutical company messages are usually weak. For example; a current mobiflex (tenoxicam - a non-steroidal anti-inflammatory drug) advert simply states "Keeping the elderly mobile" against a pleasant background of smiling elderly folk, silk and pretty flowers with someone playing with a 'slinky' toy (to symbolise flexibility and fun). The company logo is in the corner to capitalise upon any positive corporate image features. This is intended to be superficially processed. Even the number of arguments has been kept to an absolute minimum (one), although there is the required essential statutory information in small print at the bottom. Another example is evorel (transdermal oestradiol - a post-menopausal hormone replacement therapy) with the words, "Worn for effect" and supported by "The first single-membrane patch" (not a strong argument). Filling the advert is a photograph of an attractive woman in a chic jacket, and in the shadows behind her is a nude

female figure with a drug patch on her hip. Again this advert is intended for superficial processing. There are very few advertisements at the other extreme. Cytotec (misoprostil - an acid suppressing drug to counter gastric erosions from anti-inflammatory drugs) is one. The advertisement is full of health economic evidence with rhetorical questions scattered throughout and is obviously aimed at deep processing.

---

# Chapter 3

# Guidelines: Protocols, Recommendations, Policies, and Standards

ο βιοΣ βραχυζ, ηοε τεχνη μαxρη, ο οε xαιροζ οευζ, η οε πειρα σφαλεζη, η οε xρισιζ χαλεπη. Δει οε ου μονον εωυτον παρεχειν τα δεοντα ποιεοντα, αλλαxαι τον νοσεοντα, xαι τουζ παρεονταζ, xαι τα εξωθεν.

**VITA BREVIS; ARS LONGA; OCCASIO CELERIS; EXPERIMENTUM PERICULOSUM; JUDICIUM DIFFICILE. OPORTET AUTEM NON MODO SE IPSUM EXHIBERE QUAE OPORTET FACIENTEM, SED ETIAM AEGRUM, ET PRAESENTES, ET EXTERNA.**

***Life is short; art is long; opportunity fugitive; experience delusive; judgement difficult. It is the duty of the physician not only to do that which immediately belongs to him, but likewise to secure the co-operation of the sick, of those who are in attendance, and of all the external agents.***

Hippocrates, *Aphorisms*

## Summary

Clinical guidelines are an attempt to disseminate and implement the recommendations of science and expert consensus into clinical practice. Guidelines generally have a base in scientific literature review, but the most effective have also involved the intended recipients in deep processing in some way. However, it is rare for guidelines to be refined through being tested by

recipients; here lies an opportunity for making science-based guidelines more applicable in practice.

## 3.1 Clinical guidelines

David Eddy has proposed three classes of clinical policy. All are based upon scientific evidence, but are called 'standards' where there is no disagreement with regard to patients' preferences, 'guidelines' where disagreement exists, and 'options' where patients' preferences are unknown[109]. The definitions accommodate the difficulty anticipated persuading patients that the best course of action has been advocated. The definitions have little in common with the usual meanings of the words and will not be adopted in this thesis.

Where one course of action is clearly superior when the evidence is reviewed then guideline development is easy, unfortunately this is rarely the case. Either the scientific evidence is insufficient to enable a simple judgement, or there is conflicting evidence, or the evidence supports several options. Presenting the recipient of the guideline with a series of options without preference is not helpful, but how can a decision be made in these circumstances ? Weighting the options in a decision tree analysis is one approach - judgements are made by a group of experts upon which are the most important features and they are given a greater weighting - the management option with the greatest probability of achieving the objective(s) is clearly shown.[110]

Another process is Delphi analysis where groups of experts are asked to rank their preferences, these are collected and ranked, then the same experts are asked to consider the new summary ranked preferences - the process iterates until consensus is reached. The method was developed by the Rand Corporation in the 1960s as an aid to forecasting for strategic management, but has been found useful in medical decision-making.[111] Both these methods, and other forms of decision analysis, quantify judgements and make decisions explicit, but literature reviews and consensus conferences achieve their conclusions by similar exercising of judgement, though less explicitly. Where the doctors are unable to decide patients are the final arbiters - given sufficient relevant information they make their own fully informed opinion upon which option suits their needs. The trend towards greater patient empowerment however, will shift decision-making increasingly towards the patient.[112]

The most important statement in English medical jurisprudence, the 'Bolam principle',[113] states that a medical practitioner is *not* negligent if he has acted 'in accordance with the practice accepted by a responsible body of medical men skilled in that particular art'. A lecturer in law has proclaimed that the principle will be dropped in favour of comparison with written

'clinical protocols'.[114] This idea may have arisen from a confusion between legal and clinical protocols. Legal protocols necessarily have legal status, however it does not follow that medical protocols must. The courts have always recognised that there may be several acceptable ways of dealing with a medical problem. Furthermore, another legal precedent, the 'Custom test', states that three criteria must be satisfied before a doctor is found negligent. There must be a usual and accepted practice; the doctor must not have adopted that practice and; the course the doctor adopted must be one that no professional man of ordinary skill would have taken if acting with ordinary care (sic).[115] If it were possible to establish universally accepted clinical protocols only the first criterion would be satisfied. Currently, common clinical practice does this anyway. Therefore clinical protocols offer no greater threat to the profession than common practice. However, the law lecturer's pronouncement has led many authors to see 'protocols' as a threat to clinical freedom. 'Guidelines' is now preferred,[116] though in the English language neither word implies obligatory 'instruction'.[117]

In the UK guidelines represent non-punitive attempts to encourage the adoption of improved practices. The

recommendations are based on scientific evidence or expert consensus.

However, the American device, 'clinical policy' is intended to restrict clinical freedom. There are financial penalties for clinicians who deviate from a policy without sufficient justification[118].

Once established, a guideline can act as a standard against which to perform clinical audit. Other sources of standards are the common practice of a group of peers, or historic practice. Guidelines are intended to represent best practice; something to aspire to. Standards can merely be something that must be exceeded (i.e. minimum standards) or the usual activity (i.e. common practice). For these reasons guidelines and standards are not necessarily synonymous.

## 3.2 The necessity for guidelines

Guidelines are sometimes prompted by a clinical crisis,[119] [120] sometimes contradictory scientific opinions,[121] but usually they are prompted by spiralling medical costs.[122] [123] [124] The aim of clinical guidelines is to improve the quality of medical care. [125] The Department of Health understands quality as 'value for money' or 'cost-effectiveness'.

"Wide unjustifiable variations in clinical practice have been observed."[126] Guidelines are produced to reassure politicians and public that the profession has matters under control.

Thousands of guidelines have been produced worldwide. They have existed since ancient times, as in the Hippocratic *aphorisms* and Im Hotep's obstetric instructions at *Kom Ombo* in Egypt. Most of England's 200 health districts have produced guidelines. Each medical Royal College has produced or published guidelines, as have university departments, and hospitals. However, the approach is piecemeal. In the USA the policymakers tend to tackle isolated topics[127]. In the UK there is a set of guidelines, co-written by this author, which cover an entire therapeutic area of medicine (including investigations, diagnosis and treatment) in an explicit, coherent and holistic manner; 'Microbes: reference data for the management of infections in the

Wirral community'[128] This is the experimental material for this thesis.

## 3.3 Methods for developing guidelines.

The American approach according to David Eddy[129] requires ranking of options through a review of published scientific and health economic evidence. The preferred option is then incorporated into the design of a clinical policy. Work is usually commisioned by the health insurance companies (hence the emphasis on economic analysis). Evidence for the method changing clinical behaviour (in the absence of penalties) is weak.[130]

Typically British are consensus conferences[131] or working party approaches,[132133] often sponsored by pharmaceutical companies.[134135] Where the NHS has commisioned guidelines there has been an emphasis on economic analysis.[136] A group of well-informed experts gather together and debate the issues until concensus is achieved, the results are then published. The intrinsic effectiveness of this method in changing clinical behaviour appears to be insignificant[137].

Neither of these methods of guideline production are inherently effective at changing clinical behaviour. However, both do create justifiable recommendations, the American method by literature review followed by subjective preferential ranking, the British by subjective preferential ranking based upon familiar

published evidence and experience. There may be equal merit in both: the degree of scientific accuracy claimed by American authors may be overstated, particularly where epidemiological evidence[138] and meta-analyses (with compounding of errors and biases) are involved.[139] Naturally, one needs to be especially cautious about bias in guidelines sponsored by drug companies. Less apparent is the need to be sceptical of guidelines relying mainly on economic analysis.[140] It is a laudable aim to achieve the greatest benefit at the least cost. However, the only health economic analysis that approaches the sophistication of clinical decision-making is cost-utility (eg the Quality Adjusted Life-Year, QALY). This is the least developed tool in the health economist's armamentarium,[141] and has been criticised by ethicists as being inherently ageist, sexist and racist.[142]

## 3.4 Guidelines and deep processing.

Where a guideline is reported to have changed clinical behaviour a search of the report will invariably reveal some involvement of the recipents in face to face discussion. For example, in the Royal College of Radiologists' guidelines:[143] "The coordinator or a member of the radiology referral review committee also visited all the larger general practices to explain the purpose of the study, to show the participating practitioners the guideline booklet, and to obtain their approval." This is the only reference to deep processing in the whole article. Over a period of two years and involving 22 practices this guideline saved a total of £13,662. There was no reference to repetition or feedback in this study. Feedback is proven to affect clinical behaviour even in the absence of guidelines,[144] and the susceptibility of people to repeated messages is well known (especially to advertisers).[145] When examining published accounts of effective guidelines for associated deep processing those including feedback or repetition must be excluded.

In a major multi-centre study by Russell of the affect of involvement in guideline development, 84 GPs were divided into ten groups.[146] Each group developed guidelines for two of a selection of common conditions. The members of each group

were then asked to adopt two guidelines; one they had written themselves and one by another group (on another topic). The results were, "After setting their standards, (the GPs) changed their recorded practice for their condition in three ways and maintained these changes for up to two years. There was...no evidence that receiving standards set by other groups...influenced (their) practice". The Elaboration Likelihood model of persuasion is consistent with these results, the GPs had deeply processed their own subject matter, but not the others'.

Grimshaw and Russell's comprehensive review of published rigorous studies of the effect of guidelines on medical practice* divided them into two categories according to the degree of involvement of the users.[147] The guidelines they describe as *external* do not involve the users in development or dissemination. Clearly no significant deep processing is possible here. *Internal* guidelines, however, do actively involve the potential users in either development or dissemination. Elaboration is very likely to be involved here.

This author discarded all the guidelines in the review employing repetition or feedback. 51 were discarded; three external guidelines (i.e. superficial processing), and five internal

(i.e. deep processing) remained. The external guidelines produced a mean change in prescribing or clinical management behaviour of 0.17% (range 0% to 0.5%) and the internal guidelines a change of 25% (range 8% to 40%). The results are shown in figure 3.4a and are consistent with deep processing being a powerful means of achieving favourable behaviour change.

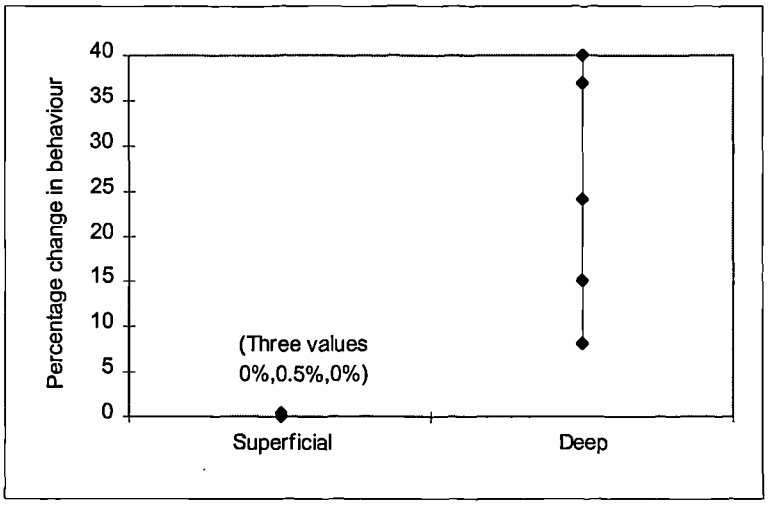



Figure 3.4a Behaviour change reported in rigorously evaluated guideline studies with or without deep processing (excluding those featuring feedback and repetition). Data from Grimshaw and Russell.[36]

* See appendix A for a critical appraisal of this paper.

# Chapter 4

# Preliminary study: local clinical guidelines

*To state even an incorrect fact boldly is to have achieved a great deal.*

Ludwig Wittgenstein (1889-1951)

## Summary

In a preliminary study a guideline developed through widespread deep processing within a district was associated with favourable attitudes to the clinical guideline's content, and behaviour consistent with it which persisted for at least 2 years.

## 4.1 Principles.

In 1991 a group of Wirral doctors and pharmacists gathered together to develop local guidelines through a *synthetic* approach. These preliminary studies established a method for guideline development and implementation which evolved into the experiment of this thesis. Some fundamental principals were adopted.[148]

Specialist and generalist clinicians constructively debating issues from their own perspectives would synthesise a more robust conclusion.*

The approach was analytical;† individual arguments were dissected and scrutinised in order to discover weaknesses. However, we must view a system as a whole before we can

---

* Georg Hegel, philosopher (1770-1831), proposed that the most effective way to make rational progress is through argument between opposing viewpoints. Triadic method requires the proposal of an opinion (thesis) by one party, and a counter proposal (antithesis) by another. Through discussion a conclusion is reached (synthesis) which is nearer to the truth than either of the original proposals This conclusion then acts as a new thesis and the process iterates, drawing closer to the truth.

† Benedict de Spinoza, cleric and philosopher (1632-1677), remarked upon the impossibility of acquiring knowledge by concentrating solely upon detail. His example; imagine a parasitic worm 'living in the blood and able to distinguish by sight the particles of the blood, lymph etc.' and able to observe the reactions of these particles with each other. If the worm confined himself to a piecemeal examination of his environment 'he would be unable to determine how all the parts are modified by the general nature of the blood and are compelled by it to adapt themselves so as to stand in a fixed relation to

satisfactorily understand the behaviour of its parts. Guideline editors and authors have to be capable of alternating between analysis and holistic appraisal. The literature of Medicine is full of sacred cows and received wisdom[149].* No fact or belief should be taken for granted. Guideline editors must be ready to challenge every assertion.† Clarification is an essential component of any analysis (see frontispiece). Criticism must be encouraged.‡

The principles of the elaboration likelihood model of persuasion were incorporated (as discussed previously). Every aspect would be audited carefully.

In summary, the principles of the approach were as follows:

---

one another' (sic). In short, a system must be viewed as a whole before we can satisfactorily understand the behaviour of its parts.

* Francis Bacon, scientist (1561-1626), summarised the origins of false knowledge as the four idols of the mind: Idols of the tribe - due to the distortion of evidence by imperfect senses (i.e. natural functional limitations). Idols of the cave - generated by the individual's prejudices, idiosyncrasies and propensities (i.e. personal biases). Idols of the marketplace - through general communications when 'ill and unfit choice of words wonderfully obstructs the understanding' (i.e. deceptions, exaggerations and inaccurate language). Idols of the theatre - dogma, mimicry and role models (i.e. the passive, unquestioning absorption of messages).

† Rene Descartes, philosopher (1591-1650), assumed an initial stance of universal doubt in his deliberations.

‡ Sir Karl Popper, philosopher (1902-1994), established refutation (i.e. searching for adverse evidence) as a much more powerful way of testing a conjecture than corroboration.

1. *Synthesis* - the interaction of opposing viewpoints.
2. *Analysis* - the scrutiny of every minute detail.
3. *Holism* - viewing details in context.
4. *Methodical doubt* - accepting no opinion, experimental evidence, or practice without question.
5. *Clarity* - to state guidance boldly and simply.
6. *Refutation* - to invite contrary opinion and counter-evidence.
7. *Mutual respect* - all contributors to be considered of equal status in order to encourage the free and frank exchange of views.
8. *Interaction* - a two-way process of development and communication.
9. *Elaboration* - the achievement of widespread and effortful scrutiny of the issues.
10. *Audit* - to examine the structure, process and outcome of guidelines.

The principles dictated that a deep processing approach be taken in evidence appraisal.

## 4.2 Wirral Medical Guidelines; part 1: Structure.

In January 1992 an editorial board comprising of a range of clinicians with a general interest in pursuing more effective healthcare was formed. They were chosen for their individual expertise:

| | |
|---|---|
| • Public Health | Dr Peter Bundred |
| • Hospital Audit | Dr John Delaney |
| • GP Audit | Dr William T Dunne |
| • FHSA* pharmacy advice | Miss Clare Dutton |
| • GP Training | Dr Murray J Freeman |
| • Continuing medical education | Dr Trevor G Gibbs |
| • FHSA medical advice | Dr Carl Onion |
| • General practice | Dr Jean Quinn |
| • Clinical Pharmacology | Dr Tom Walley |

This group were responsible for commissioning working parties to develop draft guidelines for specific topics where impact on health would be greatest.

The early working parties were as follows;

*FHSA - Family Health Services Authority, the National Health Service body which contracts general practitioner services for the patients in a district.

1. *Hypertension in the Elderly* - Dr WT Dunne (GP), Dr CJ Turnbull (Geriatrician)
2. *Menstrual Bleeding Disorders* - Dr Jean Quinn (GP), Dr Marion Jones (GP), Mr Adrian Murray (Gynaecologist)
3. *Infantile Gastro-enteritis* - Dr Peter Johnson (GP), Dr CA Bartzokas (Medical Microbiologist)

Each draft was converted into an algorithm with supporting text by the guidelines editor (myself). This was scrutinised and refined by the editorial board, and was sent to all local GPs and hospital clinical directorates for comment. Local GPs were then invited to two educational meetings per guideline to scrutinise, criticise and refine the guideline content. The meetings were structured in a round-table format and were chaired by the editor. At least one specialist was present to act as a resource for the debate. The editor modified the guideline according to the conclusions of each point in the debate. This was accomplished with the help of the editorial board who scrutinised the soundness of the arguments presented. The final draft was printed on the two sides of A4 80g paper and was laminated with plastic. A copy was then sent to every local GP and Hospital directorate. The first guideline was accompanied by a distinctive ring binder sponsored by a local firm of accountants. Sponsorship was required to help fund the project, but the most obvious source of help, the

pharmaceutical industry, were not approached for fear of introducing bias in drug choice.

There was an average (mean) attendance rate of 5.8 out of 9 (64%) at each editorial board meeting with no trend towards reduced attendance (five meetings between January 1992 and August 1992). The GP attendance rates for each meeting were as follows:

| | |
|---|---|
| Hypertension in the elderly - | 65/180 (36% of Wirral GPs) |
| Menstrual bleeding disorders - | 56/180 (31%) |
| Infantile gastro-enteritis - | 30/180 (17%) |

Attendance of editorial board members did not decline. However, quantity does not necessarily equate with quality. Some had more to contribute than others. The audit gives no indication of the quality of resource available at each editorial meeting. For example a clinical pharmacologist might be expected to have more to contribute to clinical guideline content; a public health physician more on strategy and topic prioritisation.

Attendance by the GPs at the PGEA discussion meetings declined with each topic. Was this because of declining interest in guidelines, poorly organised meetings, or less interesting topics? Too few topics have been audited to answer this. With further

guideline discussion meetings (in 1994) on different topics, if attendance rises and falls then it is likely that the topic is the cause, if decline continues then the structure is at fault. In fact most subsequent topics attracted a turnout of about 20% except depression with a turnout of 1% of general practitioners. It is likely that gastro-enteritis and depression are not attractive topics for GPs. Why this should be so when both are common and important conditions is not known, but indicates a worrying aspect of self-directed learning discussed in chapter 1 (section 1.6) above.

## 4.3 Wirral Medical Guidelines; part 2: Process

The board chose topics according to where practice was out of step with the scientific evidence (hypertension in the elderly), there was confusion (menstrual bleeding disorders), or there had been a local tragedy (infantile gastro-enteritis). Other topics that followed were led by concerns of cost-effectiveness (asthma), clinical quality (diabetes mellitus), national public health issues (coronary artery disease and stroke prevention), and local public health issues (heroin addiction). Working parties arrived at their drafts over one or two meetings. They were provided with specifications for the draft and a quiet room. Where more than two or three people were in a working party (as in some more recent guidelines) this stage was more protracted and one occasion (coronary artery disease and stroke prevention) had to be guillotined by the editor.

The early drafts arrived as rough text or diagrams. The editor made these conform to the expected format and made the recommendations more explicit. The editorial board weeded out any ambiguities and gaps. Decisions were made about which drugs were to be primarily recommended. The general principles were to:

1. use generic names
2. rank the drugs according to cost-effectiveness where there was more than one alternative
3. rank the drugs according to relative local prescribing prevalence where alternatives were of equal merit.

Comments received by post were used to refine the guideline. In practice these events were rare and generally arose from allied specialists with an interest in the topic

At the educational meetings the editor's ensured the discussion remained focused and everyone was drawn into the debate. The participants were motivated, by stressing relevance to everyday practice and the responsibility the group had to ensure that the guideline was right, to deeply process the guideline on behalf of absent colleagues. Direct individual questioning from the editor created an anticipation of being asked questions, this also ensured deep processing. Comments were recorded. Each meeting lasted one hour, was accompanied by a light lunch and was postgraduate education allowance (PGEA) approved. A revised draft was considered by the editorial board and the final version published and disseminated.

Each guideline took 27 weeks to complete as follows:

| | |
|---|---|
| • Working party (WP) develop draft | 6 weeks |
| • Refined draft discussed with WP | 2 weeks |
| • Editorial board scrutinise draft | 4 weeks |
| • Invitations to PGEA meetings | 1 week |
| • Both PGEA meetings | 5 weeks |
| • Editor refines draft | 1 week |
| • Editorial board scrutinise new draft | 4 weeks |
| • Final draft sent to original WP for final comment | 2 weeks |
| • Publish and disseminate to local GPs | 2 weeks |
| • Total | 27 weeks |

The process of evolution of the guidelines is most easily appreciated by observing the developing algorithm of the menstrual bleeding disorders guideline. Appendix 4A shows the first draft arising from the written text of the WP, the refined guideline scrutinised at the PGEA meetings, and the final published draft. The process of increasing clarity arising from the synthetic approach appears to come to life as one passes from left to right.

It was very rare for someone to avoid contributing to the debate, a typical pattern of interjections is displayed in chart 4.3a.

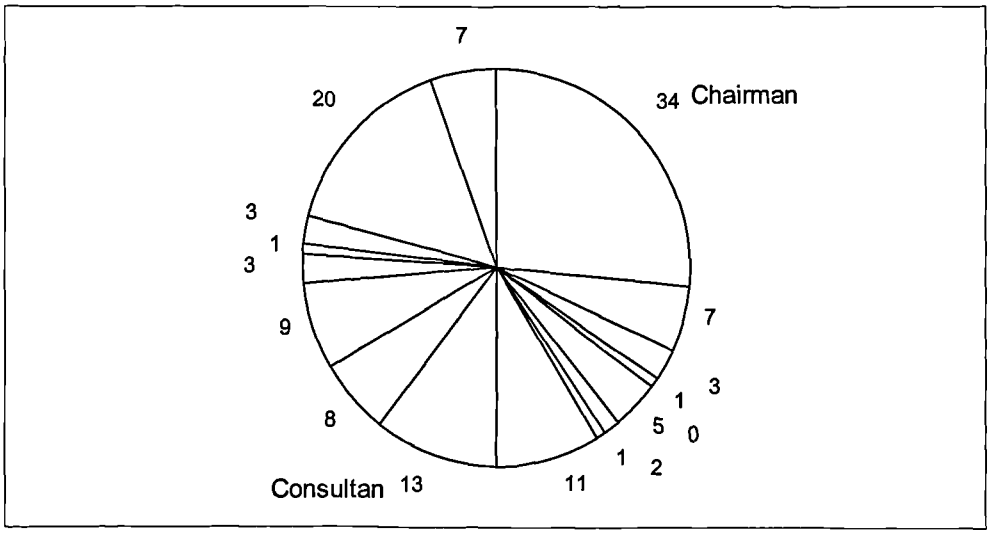



Figure 4.3a Time (minutes) each delegate was engaged in speaking at a later guideline meeting (gastro-enteritis in children). The meeting lasted one hour. Original seating positions are preserved.

Educational evaluation forms were collected after each of the events. The results for the first two guidelines are shown below. The responses are shown as agree / no view / disagree. The mode is indicated.

| Response rate: | 26%(17/65) |
|---|---|
| Overall, I found this meeting to be very useful. | 71% |
| In my opinion, this meeting was well planned. | 65% |
| The resources used in this meeting were appropriate and helpful. | 82% |
| I found the focus of this meeting relevant to my work as a GP. | 65% |
| The style of the leader was appropriate and helpful. | 82% |
| I discovered my own areas of weakness with regard to this subject. | 40% |
| I am encouraged to learn more about the subject. | 59% |

Table 4.3a Percentage agreement of participants in the deep processing meetings to the educational questionnaire.

Each guideline took 27 weeks to complete. An estimate of the total hands-on time for a guideline is:

| | |
|---|---|
| Working party | 4 hours |
| Editorial draft | 3 hours |
| Editorial board I | 2 hours |
| Redraft | 2 hours |
| GP meetings | 4 hours |
| Redraft | 2 hours |
| Editorial board II | 2 hours |
| Production | 8 hours |
| **TOTAL** | **27 hours** |

27 hours work took 27 weeks to accomplish. The greatest contributor to time elapsed was getting the working parties and editorial board actually meeting and reporting.

In summary, there was evidence of widespread deep processing at the meetings followed by favourable attitudes.

## 4.4 Wirral Medical Guidelines; part 3: Outcomes

The guideline was succinct, clear, explicit and dovetailed with the management of the condition in hospital. There was an algorithm (flow diagram) on one side of the A4 page, and supporting text on the other.

There were two main outcomes that I was interested to measure. The first was any change in attitude with regard to 'guidelines' as a concept, (from ELM* theory -see chapter 1.9 - persistent and resistant changes in behaviour could only arise from strong changes in attitude). Secondly, I wished to measure any changes in actual prescribing of the target drugs relative to expected prescribing to determine whether a persistent behaviour change had been achieved.

According to contemporary social psychology an attitude system has five components (see chapter 1.4): [150]

*ELM - the Elaboration Likelihood Model of persuasion states that attitudes arising from deep processing of an issue are persistent, resistant to counter-persuasion, and are accompanied by consistent behaviour.

1) Cognitions - What a person believes about a subject.

2) Affective responses - What a person feels about a subject.

3) Behaviour - What a person has done about a subject.

4) Behavioural intentions - What a person intends to do about a subject.

5) Attitude - the summary response; the person is either in favour of, is neutral about, or is against the subject.

Accordingly, I devised a brief questionnaire to test GPs' attitudes to 'guidelines' after the release of the guideline.* Half of Wirral GPs (90) received the questionnaire. The responses were invited on seven point Osgood semantic differential scales[151] with alternate scales reversed to prevent routine responses. The results are simplified below as positive (1 to 3) or neutral/negative (4 to 7). A response rate of 87% was achieved.

The results were as follows (yes = positive):

* See chapter 1.4 for more details of the study.

Table 4.4a Responses of Wirral general practitioners to the clinical guideline questionnaire.

| | | |
|---|---|---|
| **Cognition -** | In the treatment of many medical conditions the public are best served by a uniform approach? | yes 53% (41) |
| **Affective response -** | Does the availability of guidelines make you feel safe? | yes 60% (47) |
| **Behaviour -** | Do you often make use of currently available guidelines? | yes 59% (46) |
| **Behavioural intentions -** | Will you make more use of guidelines in the future? | yes 72% (56) |
| **Overall attitude -** | Do you think that the effect of guidelines on medical practice is beneficial? | yes 86% (67) |

I expected that attitudes would be ambivalent at best (i.e. less than 50% positive) so the results were promising. Overall attitude to guidelines was strongly positive (86%) as was the intention to follow them (72%). These results are similar to those obtained through an attitude questionnaire sent to Dutch GPs by the Nederlands Huisartsen Genootschap (NHG - the Dutch college of General Practitioners) which sets national clinical standards for Holland;[152] 80% of their respondents were in favour of national clinical standards.

The first exemplar guideline on 'hypertension in the elderly' was developed by the method described and principally recommended bendrofluazide 2.5mg once daily. The main outcome measured was the differences in prescribed daily doses (PDD) of bendrofluazide 2.5mg tablets per quarter per 1000

prescribing units ('PU' - everyone under 65 counts as one PU, and everyone 65 and over as 3 PUs) between the intervention district and England as a whole.

Comparison of the intervention district (Wirral) with England data demonstrated a median difference of 122.49 PDD before and 206.34 PDD after guideline production, this change is statistically highly significant (Mann-Whitney two-tailed $p < 0.0001$; 95%CL = 36.51 to 104.77).

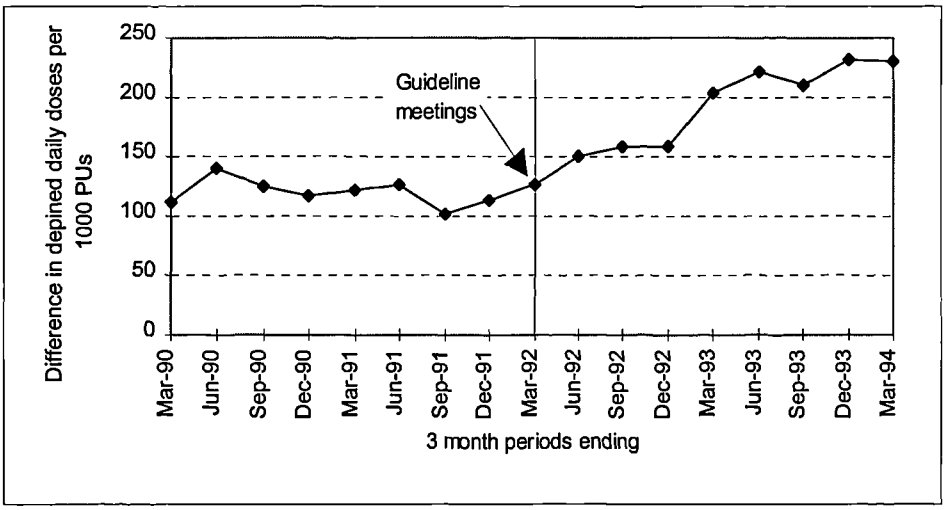



Figure 4.4a Difference in number of prescribed daily doses of bendrofluazide 2.5mg per 1000 prescribing units (age-weighted population) between the district and England as a whole.

Grouped regression analysis showed no significant difference (0.89) in slope gradients before guideline production ($p = 0.35$, 95% CI = -3.97 to 5.76), but the difference in slope gradients after (12.95) is statistically highly significant ($p < 0.0001$;

95% CI = 8.17 to 17.73). This suggests that the change in clinical behaviour persisted for at least two years.

An objective of the guideline project was to produce strong positive attitudes through interactive education. This was then hoped to result in consistent behaviour. Attitudes formed through direct experience with the attitude object or issue are more predictive of behaviour than those formed indirectly.[153] The preliminary study indicated that the method should taken further. Several similar guidelines on other disparate topics have followed. However, there was scope to refine the method, cover wider topics, and to conduct a controlled scientific investigation.

This thesis will test the fundamental method (described above) of development and dissemination of a set of guidelines in a controlled experiment.

## 4.5 Infection management as the test model.

The chosen therapeutic area for testing the deep processing model for guideline development and dissemination in general practice was infection. According to symptom diaries, adults experience a mean of 3.9 symptoms per fortnight.[154] Many of these symptoms suggest infection:

| | |
|---|---|
| 12% will experience..... | sore throat |
| 32%..... | cough or catarrh |
| 18%..... | cold, 'flu, or runny nose |
| 3%..... | diarrhoea |
| 4%..... | sores or ulcers |
| 2%..... | a temperature. |

The rate of reported symptoms seem high, but very minor symptoms were recorded (e.g. aches, abdominal discomfort etc.). One hundred and eighty-one people per thousand population will consult their GP with an infection or infestation per year.[155] The typical annual prevalence in a general practice population of 2500 is:

| | |
|---|---|
| Upper respiratory infections | 600 |
| Acute tonsillitis | 100 |
| Acute bronchitis | 100 |
| Acute otitis media | 75 |
| Acute urinary infection | 50 |
| Vaginal discharge | 30 |
| Pneumonia | 20 |

So infection is a common cause of morbidity in the community. Many episodes represent minor self-limiting conditions. It is clear that more people experience sore throats than consult their doctor for a sore throat; either GPs are seeing the more severe cases or those unable to cope (or the symptomatic data is incorrect). In the first case it is imperative that the condition is treated effectively by the GP, in the second case it is important that doctors and other healthworkers give out consistent and effective educational messages to the public. To do this they must first decide upon a common approach to sore throats, the current variation confuses patients and undermines the work of colleagues. Infection is a field of medicine where guidelines should be highly beneficial and was therefore chosen to test the deep processing approach.

# Chapter 5

# Purpose and aims of study: two levels of processing

> *The decision to reject one paradigm for another is always simultaneously the decision to accept another, and the judgement leading to that decision involves the comparison of both paradigms with nature and with each other.*
>
> Thomas S Kuhn. 1962 [156]

## Summary

The main study sought to test scientifically whether depth of processing in intended recipients of a set guidelines affected subsequent relevant knowledge, attitudes, behaviour and patient outcome. The guidelines covered a complete field of medicine (infectious disease in the community) in order to demonstrate that an holistic and more efficient approach than the current piecemeal one was feasible. The experiment polarised interventions on two groups of practices into either deep processing interviews (visits) or superficial processing events (lectures). In both cases educators promoted the same educational messages. A control

group established the baseline against which to measure relative performances.

## 5.1 Hypothetical matters.

*The hypothesis.*

If general practitioners deeply processed* the messages in a set of scientific guidelines consistent and persistent changes in clinical behaviour would follow.

*The sub-hypothesis.*

Deep processing in guideline development would result in attitude changes more persistent, resistant, and consistent with behaviour than superficial processing†.

*Refutation.*

As advocated by Popper, the experiment attempted to refute the above hypothesis.[157] By Popperian standards the hypothesis was a good one in that by implying measurable changes in actual clinical behaviour the hypothesis exposed itself readily to testing. If it could be shown that changes in behaviour were absent or did not persist following deep processing the hypothesis would be disproved. Similarly, if deep processing did not result in more

*To think about an issue and its supporting arguments in an effortful and critical manner.
†To rely upon simple sensory information to assess an issue.

change in behaviour than superficial processing the sub-hypothesis would be disproved.

*Logic.*

In a black and white world the hypothesis could have been tested by comparison of single cases. The Popperian conjecture and refutation model of establishing scientific validity is based upon the logical rule *modus tollens*.[158] This establishes that if an event p is stated to result in a consequence q, and there is no q then there should be no p either. Symbolically:

1. Hypothesis If (p), then (q).
2. Corollary Either: (not p) or (q)
3. Observation (not q)
4. Implication Therefore (not p)

*Quod erat demonstrandum.*

For example; 'fire causes smoke' - so there is either no fire or there is smoke, therefore, if there is no smoke there is no fire: (note that this is not the same as 'there is no smoke without fire' which assumes that *only* fire can cause smoke - a familiar deviation from logic amongst gossips!). However, if events (p) and (not q) could be shown to coexist then the hypothesis must have been false:

| | |
|---|---|
| 1. Hypothesis | If (p), then (q). |
| 2. 1st observation | (not q) |
| 3. 2nd observation | (p) |
| 4. *Modus tollens* 1) 2) | (not p) |
| *5.* Addition of 3) and 4) | (p) and (not p) |

*Reductio ad absurdum.*

The hypothesis must be wrong because the state of coexistence of (p) and (not p) is impossible. For example; 'fire causes smoke', there is no smoke, so there can be no fire, but there *is* fire, so fire has not caused smoke, the hypothesis must be wrong. In a perfect world therefore, a solitary case that did not comply with an hypothesis would utterly refute it.

In reality things are never so clean and simple, however. The accuracy of measurement is unlikely to be 100%, unseen chance factors would also be creating or destroying occasional events, and some data would be lost. In these circumstances reliance on single events would be a lottery. To reduce the chances of falsely accepting or refuting the hypothesis, the long-run frequencies were examined. A chance association of p and q could occur once or twice by chance, but were unlikely to be repeated dozens of times in succession. This is equally true of a chance non-association of p and q.

## 5.2 The purpose of the experiment.

'Deep processing is superior to superficial at changing clinical behaviour'. This experiment attempted to test that statement by comparing the effect of these two approaches. General practitioners were exposed to a set of guidelines advocating best clinical practice through either superficial or deep approaches. Both would be compared with a state of not being exposed to either approach. The purpose then was to discover firstly whether the deep approach worked, and secondly whether it was more effective than a standard superficial approach.

## 5.3 Aims.

*1) To develop a set of guidelines advocating best practice in a field of medicine.*

The guidelines had to justifiable on current scientific evidence rather than common practice because practice is slow to follow science.[159] Therefore, if the guidelines were followed an observable change in behaviour should have resulted. They should also have been explicit and unequivocal for two reasons. Firstly, they must have boldly invited criticism for the purposes of deep processing. Secondly, consequent changes in behaviour would have been be easier to measure if a limited number of options were advocated.

Infectious disease was chosen as the field of choice for the experimental guidelines because there was evidence of variation in its management, it overlapped every other field of medicine, prescribing and special investigations data was routinely collected by the Prescription Pricing Authority and the local microbiology laboratory.

*2) To expose general practitioners to either superficial or deep processing of the guidelines, but not both.*

One group would receive the guidelines and be informed about them via a standard superficial method with little opportunity for deep processing (*superficial group*). The other would be exposed to attempts to facilitate deep processing about the guideline content (*deep group*). A third (*control*) group would not be exposed to the guidelines at all and their behaviour would act as a benchmark by which to measure the two intervention groups.

For meaningful comparison the superficial method would have to be realistic and familiar. Didactic lectures are a traditional, popular, and common way to present clinical information to general practitioners.[160] Therefore the superficial group would be exposed to passive receipt of the guidelines and would be invited to attend a didactic clinical lecture event on the guideline content.

The deep approach must also be realistic and familiar, or it would not be reproducible. Outreach visits to medical practitioners in their surgeries by other professionals were an established method of engaging practitioners in thinking about therapeutic issues.[161] Medical advisers were employed by Family Health Services Authorities (FHSA) in England to perform this function.[162] The deep group were therefore visited by the FHSA medical adviser and the local specialist in the chosen field (consultant

medical microbiologist). During the visit engagement of the general practitioners in deep processing of the guideline content was attempted.

The control group would not receive the guidelines and would be isolated from the superficial and deep processing interventions.

*3)To identify and measure variables which would indicate a change in behaviour.*

Where practice was already consistent with the guidelines no change would be detectable no matter how effective the approach. The most fruitful areas to observe any changes would be where common practice differed most from the advocated course of action. Activities that differ most from those expected would be identified. In some instances this would be a change in the number of cases investigated or treated in a particular way, In other cases it would be a change in proportion of advocated activity compared to non-advocated activity, for example an increased ratio of penicillins to cephalosporins use overall. In still other cases the use of strongly advocated, but currently seldom used investigations or drugs, or of activities strongly advised

against would be applicable. In some instances it may be possible to identify measures of favourable patient outcome.

The choice of indicators would be based upon evidence of current variation from optimum practice before data analysis. Retrospective selection only of indicators that support the hypothesis would invalidate the experiment. Having said that, it might have been impossible in advance to know what the change would be, merely that there should be a change. For instance, the guidelines covered some aspects of the use of mid-stream specimens of urine (MSSU). We might have expected a change in the rate of referral of MSSUs to the laboratory. Greater awareness of MSSU issues could have increased the use of MSSUs, conversely greater confidence in the advocated blind treatments could have reduced the use of MSSUs. All that could be anticipated was a difference, whether it would be positive or negative could not be stated in advance.

Indicators of clinical behaviour in the three groups would be measured prospectively and compared. Data on prescribing activity and laboratory investigation rates were already collected routinely. It would therefore be possible to examine and compare some of the variables both longitudinally and cross-sectionally.

## 5.4 Design

*1) To identify three comparable groups of general practitioners.*

Three samples of general practitioners would be drawn from the same geographical area so that they would be exposed to the same external influences. They must also have had the same overall distribution of characteristics known to affect clinical behaviour.

*2) To control the experiment.*

The three groups would be protected against unbalanced external influences that were within my sphere of control. No other specific work on infections targeted at any general practitioner would be undertaken by the FHSA or hospital, including clinical audit in the field of infection. The subjects would be 'blinded' as much as possible, they would be unaware that they were being especially monitored or this could have affected their behaviour (Hawthorne effect* ).[163] Similarly the investigators would be blinded, as far as the data goes, for the period of the experiment,

*Hawthorne effect - first described in the Relay Assembly Test Room experiment at the Hawthorne plant of the Western Electric Company in Chicago 1927-32. Employee performance improved simply because of the novelty of being involved, and the extra attention they received.

or they could otherwise subconsciously apply remedial action to non-conforming practices.

The educational messages behind both the superficial and deep approaches, and the amount of time the subjects are exposed would be identical. The only difference would be the depth of processing, otherwise any effect could have been due to different emphases or time applied. However, if the same educators were used to deliver both the lectures and to carry out the visits any effect could be due to their being better skilled at one approach than the other. The superficial approach would be given its best chance to succeed and experienced lecturers would deliver the lectures. Conversely, to be realistic, the local medical adviser and consultant would carry out the visits. As they were briefed in deep processing techniques they were likely to perform to a standard comparable to the professional lecturers.

'Contamination', the spillage of intervention or effect, between groups was prevented. The control group did not have access to the guidelines throughout the study period. The superficial group would not have opportunity to deeply process the information. The deep group would not be allowed to assimilate the information entirely in a passive superficial way.

*3) To motivate the intervention groups to partake in the educational events.*

The superficial group would be motivated to read the guidelines and attend the lecture event by superficial methods (extrinsic motivators).[164] The guideline book was attractive, from 'experts' and produced by the Health Authority. A free buffet lunch and postgraduate education accreditation points (P.G.E.A.) and an impressive double bill of respected lecturers enticed GP member of the superficial group to attend the launch event.

Superficial, (extrinsic) motivation would not be sufficient for the deep group. They would have to be motivated to engage in effortful deep processing of the messages. That practitioners were obliged to discuss their prescribing with the FHSA medical adviser by law, and that the visits could attract a fee acted as extrinsic motivators to gain access to them. However, to ensure that they engaged in deep processing required 'intrinsic' motivation. Intrinsic motivation describes a state of personal conviction and commitment to an issue. For an external agent to motivate a subject intrinsically there must be an alignment of the perceived potential benefits of involvement in the issue with the subject's personal aims.

*4) To develop expert- and specialist-driven guidelines on infection (the first edition).*

One way to develop valid clinical guidelines would be to capture the local general practitioners' opinions on best practice. This would be legally defensible under the Bolam principle in that management would be advocated which an ordinary practitioner of ordinary competence would pursue.[165] However, if such guidelines were followed assiduously limited improvement in clinical practice is likely as practitioners would merely be aspiring to do what they already aspire to do. Not only would this achieve little benefit to patients, but changes in behaviour were likely to be too marginal to observe.

What was required were guidelines that, where appropriate, advocated behaviour that was significantly different from current practice. The only time that departure from 'accepted' practice would be legally acceptable would be when the departure could be justified on good scientific evidence or unique clinical grounds.[166] The guidelines would therefore be based upon good scientific evidence rather than current practice. In this way patient care would be improved if the guidelines are followed, and changes in behaviour measured more easily.

*5) To improve the guidelines following suggestions from the recipient general practitioners (the second edition).*

The value of the views of guideline recipients had not been recognised. If users' views were solicited not only should they have felt more involved in the process, but any impracticalities in the science-based guidance would have been identified.[167] The first edition represented a scientifically justifiable way to manage infections. As much of the evidence was be derived from controlled trials or where diagnosis is supported by comprehensive laboratory investigation it was not likely to be entirely applicable in general practice. In practice many conditions present before they are florid, the working diagnosis then represents a best guess and options need to be kept open. Controlled scientific trials usually exclude the complicated cases common in everyday practice and still requiring treatment.

The practitioners in the deep group would be invited to scrutinise the guidelines, and therefore be likely to spot impracticalities. These observations and subsequent suggestions would be collected during the study period for incorporation into a more practical second edition at the end. Therefore, practitioners would feel more involved in and responsible for the project, and their collected observations would give an impression of the

magnitude of contribution from actual practitioners necessary for guidelines to become robust in practice.

*6) To measure the structure, process and outcomes of the project.*

It was not to be taken for granted that deep processing would automatically follow discussions. It would be necessary to demonstrate a significantly different level of processing in the two intervention groups. Some measure of the degree of deep processing in the deep group relative to the superficial group would be undertaken. To this end the lecture event (superficial group) and practice visits (deep group) would be recorded on videotape with particular reference to audience reactions. Independent experienced psychologists would then score individual practitioners according to indicators of deep processing.

Any absence of increased deep processing in the deep group could be due to the models being ineffective in this setting or, and more likely, the models not being followed faithfully. Adherence to a study protocol would negate the latter. Records of the interventions would record the degree of successful adherence.

A false impression of the effectiveness of the models could be gained if any additional interventions were applied during the study. For example, rewards for following the guidelines or remedial attention to poor. Avoidance of contact with the practices in association with any infection management issue except as in the study protocol would be avoided.

The immediate measurable outcomes of most presentations of infection in general practice were investigations and prescriptions of antimicrobial drugs. Both these sets of data were collected routinely and systematically by the local microbiology laboratory, and the Prescription Pricing Authority respectively. Both sets of data were based on general practices and were therefore compatible.

For at least one specific infection a detailed clinical study would be performed. This would examine the rate of investigation per case, treatment use, and recurrence rates. This audit would be carried out at the end of the study period to examine the predicted long-term effects on practice. If the audit was repeated after feedback of the initial audit results it would be possible to compare the effect of clinical audit on changing clinical behaviour with the deep processing approach in the project.

## 5.5 Confounding factors

The experimental hypothesis was a specific application of the 'Elaboration Likelihood Model' of Petty and Cacioppo. Their hypothesis is that if strong messages are deeply processed resulting attitudes are persistent, resistant to counter-persuasion, and followed by consistent behaviour. One problem is that any favourable changes in behaviour that we observe might not have been due to attitude change. Falsely concluding that attitudes had changed would undermine the theoretical validity of this and subsequent work. It was therefore essential that a parallel study of attitudes to the guideline topics before and after the main study was undertaken. If there were changes in behaviour in the deep processing group then there should have been associated changes in attitude.

## 5.6 Message strength.

The definition of a strong message in the Elaboration Likelihood Model was unfortunately somewhat circular. A strong message was one that, when deeply processed, results in a favourable attitude. And a favourable attitude was one that resulted from the deep processing of a strong message. The guideline messages were assumed to be strong because they had arisen from scientific fact and reasoning. The definition needed to reflect this specifically and be compatible with the definition above. *A strong scientific message was one that had supporting facts and sound logical arguments so that the deeper it was processed the more convincing it became.* This differentiated it from strong unscientific messages based upon irrational arguments that were still compelling when deeply processed. For example:

*'Trimethoprim is the drug of choice in uncomplicated cystitis'*[168] is a strong scientific message, because:

1. local data shows that 70% of cystitis infections in the local community are susceptible to it.[169]
2. 50% of cases resolve spontaneously anyway.[170]
3. It is less prone to side-effects than its usual alternative, co-trimoxazole.[171][172]
4. Higher cure rates at the expense of increased risk of side-effects or more costly drugs were not worth pursuing as the risks of serious sequelae from uncomplicated non-recurrent cystitis are infinitesimal.[173]

*'Co-trimoxazole is the drug of choice in cystitis'* is a strong unscientific message because:

1. The leading brands are made in this country.
2. The syrup form usually had a more pleasant taste than trimethoprim syrup.
3. Two antibiotics must be better than one.
4. Urinary tract infections admitted to hospital were usually resistant to trimethoprim alone.

Both types have compelling supporting arguments when scrutinised. 70% susceptibility was sufficient in a low risk infection with a high spontaneous resolution rate (scientific message 2). However, it is an important social duty to support home industry (unscientific message 1). The difference was that the scientific arguments can be supported and evaluated by reference to current scientific evidence, the unscientific cannot. Note that a strong argument was not irrefutable; for example 70% of cystitis infections would be sensitive to trimethoprim, but 'sensitivity *in vitro'* was not the same as 'effectiveness *in vivo'* which is what counted,[174] and whether 70% sensitivity was sufficient was a matter of judgement on the behalf of the practitioners. Strong messages therefore stimulated a deeper and more scientific level of scrutiny and discussion.

Another problem was the polarisation of superficial (didactic lecture) and deep (practice visit) approaches. Although didactic

lectures are still common, their shortcomings are known to medical teachers. The university lecturers involved in the superficial presentations would have to restrain their urge to incorporate deep processing features. A strict protocol for the lectures including the messages to be included would have to be set beforehand and adhered to. On the other hand if the didactic lecture was too stereotypical it would not be a reasonable comparator. The standard approach was to allow a few minutes for questions at the end of lectures.

To be sure that like was compared with like, the key messages and topics covered in the visits should be identical to those in the lectures. However, deep processing encourages the subject to dictate the direction of discussion. However, to err on the side of caution the agenda would be observed as strictly as possible; otherwise, control of educational factors would be lost with the consequent risk of making type 1 (falsely rejecting the null hypothesis) or type 2 (falsely accepting the null hypothesis) errors. Recordings of the lectures and a sample of the visits would allow verification later.

Another concern about the visits was that the Elaboration Likelihood Model described first the necessity of engaging a person on a superficial level before being able to persuade them

to perform deep processing. The deep approach was therefore in fact a combination of an initial superficial approach leading to deep processing later.

The Ford motivation method incorporated in the visits relies upon successfully anticipating the subjects' goals in life. The work of Bradley on critical event analysis allowed us to infer which important personal goals might have been obstructed (see chapter 1.5).[175] However, in view of the uniquely personal nature of goals and in the absence of any other published work on doctors' goals we would largely rely on intelligent guesswork. The visitors would have to watch for expressions of interest and disinterest as potential goal areas were raised in each visit.

The Elaboration Likelihood Model is a potent persuader. Some commercial advertisers, politicians, and military organisations had successfully used it and its antecedents to persuade people to buy, vote, and fight against their own best interests. It was imperative that medical educators only applied it in practitioners' best interests. By being completely open about the model two things were achieved. The practitioners would see that there was no hidden agenda. They would learn how deep processing works and was equipped to control subsequent attempts at counter-persuasion from commercial sources.

Commercial organisations adopted a covert stance, this would be unethical in education where there was an implicit contract of trust.

It was anticipated that when practices were visited it would not be uncommon for one or two partners to be absent, so there would be a dilution of the intervention. Attendance at lecture events is usually even poorer so vigorous attempts would be made to raise as big an audience as possible at the lecture event. In this way no subsequent differences could be explained away purely on the basis of poor turnout. Visits would similarly not exceed the length of time that the superficial group were exposed to their lectures. Otherwise differences in effect could be due to differences in exposure time.

A balance had to be struck as to how extreme the scientific approach in the first edition of the guidelines should be. They should be based on expert and scientific evidence and opinion so that they could be robustly defended in the deep processing visits, but too much would make it look as though they had been written by academic idealists isolated from real clinical medicine. If credibility was lost at first impression in this way then the attempt to engage the practitioners in deep processing would fail at the superficial stage. The strategy was to offer the first edition as a

first draft and invite constructive criticism. A potentially negative feature was then turned positive because the practitioners were likely to offer valuable suggestions for improvement.

## 5.7 Anticipation of problems with data and analysis.

*1) Variation with time.*

Infections could be unpredictable, some more so than others (table 5.7a).[176] For any particular infection the number of cases in a month would vary from month to month, and could even differ in the same months in adjacent years. Most minor infections did not lead to a medical consultation. Therefore the cases recorded represented only the margin and were susceptible to considerable swings in response to social factors. For example people with mild infections might ensure that they consulted a doctor earlier before bank holidays when medical services might be difficult to access, or after media scares about, for example, meningitis or whooping cough.

| Infection | Weekly incidence range (per 100,000 pop.) | Approximate year on year trend |
|---|---|---|
| Acute bronchitis | 50 - 150 | + 3% |
| Influenza | 4 - 90 | na |
| Common cold | 70 - 250 | + 3% |
| Acute otitis media | 35 - 110 | + 2% |
| Acute tonsillitis | 65 - 110 | - 1% |

Table 5.7a Annual range of mean weekly incidence of respiratory infections, and an approximation of the underlying annual trend - from data for England and Wales over 10 years (1980-89).

*2) Variation between practices.*

A survey of the prescribing habits of general practitioners in the Wallasey locality of Wirral showed tremendous variation in the volume and range of antibiotics prescribed.

| Drug | Mean | Standard deviation |
|---|---|---|
| Broad spectrum penicillins | 83.8 | 29.6 |
| Macrolides | 39.2* | 41.2 |
| Cephalosporins | 34.4 | 28.6 |
| Tetracyclines | 27.9 | 21.4 |
| Co-trimoxazole | 15.6 | 12.4 |
| Penicillin V & Benzylpenicillin | 12.6 | 10.1 |
| Quinolones | 11.6* | 21.6 |
| Trimethoprim | 9.4 | 6.8 |
| Penicillinase resistant penicillins | 7.4 | 7.2 |
| Nitrofurantoin | 1.3 | 1.8 |
| Total | 243.2 | - |

Table 5a) The mean prescribed items per 1000 population for general practices in Wallasey, Wirral in the months February, March and April 1992. (Population 60,447 served by 18 practices). NB the mean occasionally exaggerates the rate (marked *) because of outliers.

The actual lengths of the courses prescribed varied. There was variation in the rate of laboratory investigation. For example some practices investigated every case of cystitis at first presentation, others virtually never did. Some practices withheld antibiotics for minor virus infections more than others, relying more on simple advice and symptomatic remedies.[177] Outliers (i.e. unusually extreme individuals) have a significant effect on the mean value.[178] We would be unable to examine within practice variation as the prescribing data was not reliably linked to

individuals. Practices have exclusive prescription pad numbers, but partners within the practices do not.

The study examined differences between practices, and before and after the intervention regarding prescribing and investigation frequencies for infection (see chapter 7). This would be complicated by the variations described above. Our aim was to reduce the variation as much as possible in data collation before analysis proceeds.

Aggregation of data would help to reduce variation. For instance monthly data could usefully be added up into quarterly (3 monthly) data. This would help to iron out month to month variation. This data could also be added up into 12 months. This would resolve seasonal variations. Between-practice variation could be reduced by aggregating the data on all practices in each experimental group. However, this would tend to reduce statistical analysis to the relatively insensitive study of proportions. Rolling averages would be unhelpful as seasonal peaks were still highly variable; also data collection would extend the study period well beyond the 12 months designated with increasing difficulties in maintaining control.

About 40% of the variation in general practitioner prescribing has been associated with situational factors. Forster and Frost

showed at Family Practitioner Committee level that the local Standardised Mortality Rate (SMR), the supply of doctors per head of population, and the proportion of elderly female patients significantly affect the volume and cost of prescribing.[179] SMRs were not available at practice level, but were available at aggregated electoral ward levels where significant geographical variation was evident. For this reason a geographical balance in the three experimental groups was attempted. At local level the supply of doctors was affected by the availability of (supernumerary) trainees. The infections section of the NHS prescribing data was the least sensitive to population age and gender so allowance for these factors was unnecessary (D Lloyd, Prescribing Research Unit, Leeds - personal communication). Financial incentives were very effective at changing prescribing practice, all Wirral fundholding GPs (who were able to invest any prescribing savings in their practices) had reduced their prescribing cost relative to the district average since becoming fundholders. There would have to be a balance of all these factors across the three groups.

| **Factor** | ***Proxy at practice level*** |
|---|---|
| SMR | *Geographical location* |
| GP supply rate | *Trainee* |
| Financial Incentives | *Fundholding practice* |

Other likely confounding factors were size of practice, and previous behaviour. With several factors stratification and randomisation* would have been exceedingly complex. A more appropriate method of matching the groups by minimisation† would be undertaken.[180] Practices would be assigned to groups to minimise the overall imbalance of known or likely major confounding factors for prescribing. To achieve unbiased allocation, this would be performed by an independent researcher aware only of the practice characteristics and blinded to the identity of the practices

Standardisation of the data in terms of frequency per head of population would reduce variation due to registered list size. The above would help to reduce between-practice variation and would facilitate cross-sectional analysis. That is, examining differences between the three groups of practices at the same points in time. The matching would not be accurate enough to employ paired tests.[181]

The perfect match would be identity, or A ≡ A. In other words the best match for each group was itself. Provided that there were

---

*Dividing a population into groups according to important likely confounding variables and randomly selecting a proportion pro rata from each division.
†Random assignment of individuals to groups with subsequent reassignment of a few individuals to achieve the best balance of important confounding variables.

no other significant changes in a practice then the difference in relevant activity before and after the intervention could be analysed, a longitudinal analysis.

Statistical power is the ability of the data in the study to show any differences and was most heavily influenced by the number of events upon which data was collected.[182] Our study was necessarily confined to one district because this was the maximum area over which we could reliably control potential confounding factors. To maximise statistical power all 69 practices in the district had to be incorporated in the experiment if possible. One way of controlling for confounding factors was to allocate subjects to groups entirely randomly, if the population was large enough then the factors would become increasingly equally distributed. Unfortunately 69 subjects (a maximum of 23 in each group) would be inadequate for this purpose; it was very likely that one or more significant confounding factors (such as fundholding) would be over-represented in a group. Hence the choice of minimisation as the allocation method. It was intended that minimisation would distribute the known confounding factors equally between groups.

As identified earlier, it would necessary to predict the activities where significant change was possible. In the absence of

data linked to individual cases it would be difficult to predict this in advance without a major piece of clinical audit. In lieu of this, epidemiological modelling allowed me to predict what proportions of antibiotics would be prescribed for the expected pattern and volume of infection if the guidelines were slavishly followed. This could then be compared with actual prescribing and the greatest variations from the guidelines would be exposed.

I predicted incidences of infection on Wirral for the quarter ending October 1993 from the third national morbidity survey data.[183] When the 'Microbes' guidelines were applied stringently to these expected cases the model generated a profile of the expected pattern of antibiotic prescribing. This can then be compared to the actual prescribing as collected in PACT data. Rare infections (less than 1% of expected) are not included in the analysis:

| Treatment | Prescriptions expected from guidelines | Percentage expected % | | Prescriptions actually dispensed | Percentage prescribed % |
|---|---|---|---|---|---|
| No antibiotic | 29822 | N/A | | N/A | N/A |
| Co-amoxiclav | 5023 | 36.5 | | 9046 | 14.8 |
| Amoxycillin | 3610 | 26.2 | | 19902 | 32.5 |
| Trimethoprim | 2093 | 15.2 | | 5406 | 8.8 |
| Penicillin V | 1413 | 10.3 | | 6015 | 9.8 |
| Erythromycin | 1116 | 8.1 | | 3830 | 6.2 |
| Cephalexin | 523 | 3.7 | | 4071 | 6.6 |
| Cefaclor | 0 | 0 | | 3004 | 4.9 |
| Cephradine | 0 | 0 | | 501 | 0.8 |
| Cefixime | 0 | 0 | | 990 | 1.6 |
| Ampicillin | 0 | 0 | | 644 | 1.1 |
| Tetracyclines | 0 | 0 | | 1111 | 1.8 |
| Co-trimoxazole | 0 | 0 | | 3173 | 5.2 |
| Clarithromycin | 0 | 0 | | 916 | 1.5 |
| Ciprofloxacin | 0 | 0 | | 1623 | 2.6 |
| Pivampicillin | 0 | 0 | | 10 | 0.1 |
| Ofloxacin | 0 | 0 | | 426 | 0.7 |
| Nitrofurantoin | 0 | 0 | | 546 | 0.9 |
| Norfloxacin | 0 | 0 | | 80 | 0.1 |
| Total | 43600 | 100% | | 61294 | 100% |

Table 5b The number of prescriptions for each antibacterial drug in Wirral in the months August, September, October 1993 (guideline drugs shaded). The number expected from epidemiological modelling and recommended drugs in the guidelines were also shown. Note that the model significantly under-estimates antibacterial prescribing. (The percentage expected figures were all ± 1% as rare diseases [≤ 1% of cases] have not been incorporated in the model)

Discrepancies in the volume of prescriptions suggests that there might be a difference in the local morbidity (or prescribers) relative to the national picture. Overall, Wirral GPs were high volume prescribers, perhaps reflecting the district's high levels of social deprivation. Also, the GPs may have been following their own , or other, guidelines rather than the ones under test. Notwithstanding this the model generated some potential

indicators of adherence of the study guidelines. The notable differences between the guidelines 'ideal' practice and real practice were as follows:

1. The guideline drugs were only used in 78.8% of cases; the guidelines advocated 100%.
2. Co-amoxiclav was less of a proportion of prescriptions than advocated.
3. The ratio of trimethoprim to co-trimoxazole was 1.7; the guidelines advocated no co-trimoxazole.
4. The guidelines advocated more cephalexin as a proportion of total cephalosporins prescribed.
5. 1.8% of antibiotic prescriptions were for a tetracycline; the guidelines advocated none.
6. The newer antibiotics accounted for 12.3% of antibiotics; the guidelines advocated none.

To refute the null hypothesis it would be necessary to show significant changes in these proportions of prescriptions after the intervention.

Different practitioners were likely to prefer different learning styles. The deep group might respond simply because it contains a preponderance of people who were fascinated by new things, like guidelines for instance. It would be necessary to demonstrate that learning styles were well distributed in the three groups. This

would be augmented by an auxiliary analysis of how each type tends to respond to the lectures or visits. The standard learning style questionnaire (Honey and Mumford 1986) would be sent to each participating practitioner during the study with feedback to them after completion of the experiment.

The use of the practice as the basic unit presents a problem in that the performance of the practice was a function of the mix of partners. In some cases the variation within the practice may be great with one or more outliers. In these circumstances an important change in the moderate partners may not be observable if there was no change in the outliers' performance. This would dilute the ability of the data to detect differences.

The adoption of a three-arm trial (control / superficial / deep) instead of a two-arm trial (superficial / deep) reduces the statistical power of the study even if all 70 Wirral practices take part, but tests two hypotheses, firstly that deep processing is effective, and secondly that it is more effective than superficial. This study was analogous to a phase III drug trial.[184] I had established that the deep processing approach does not have an adverse effect on ` learning facts (phase I) - *see chapter 1.7.*[185] I had also

---

*Cefaclor, Clarithromycin, Cephradine, Cefixime, Ciprofloxacin, Pivampicillin, Ofloxacin, Norfloxacin.

demonstrated that it was effective in altering behaviour (phase II) - *see chapter 4*.[186] Deep processing had been followed by favourable behaviour change. However, effectiveness of deep processing relative to superficial processing was unknown (phase III). If the effectiveness of the standard model (superficial evaluation, passive receipt of information, and didactic lectures) had been shown then it would only be necessary to compare the two methods. However, it was not known whether the standard method was ineffective or even adverse. In these circumstances it was necessary to include a control group. It was possible that superficial processors could fare better or worse than control or deep processors.

# Chapter 6

# Trial protocol, method and progress

**Zweimal zwei gleich vier ist Wahrheit,**
**Schade, daß sie leicht und leer ist,**
**Denn ich wollte lieber Klarheit**
**Über das, was voll und schwer ist.**

*Twice two equals four 'tis true,*
*But too empty and too trite.*
*What I look for is a clue*
*To some matters not so light.*[187]

Wilhelm Busch 1909

## Summary


*Time period*

Two years: November 1992 to November 1994 inclusive.

*Subjects*

All (69) general practices in an English Family Health Services Authority district comprising 180 general practitioners and one microbiology laboratory serving a population of 350,000.

*Setting*

The Wirral peninsula, a district in north-west England separated from Liverpool by the River Mersey, from Wales by the River Dee and from Chester by a rural belt. Unemployment and other indicators of social deprivation are pronounced in some localities. The main town is Birkenhead.

*Experiment*

To compare the effectiveness of deep and superficial processing methods of guideline development and implementation in the management of infections in general practice.

*Main measurements*

Verbal evidence of deep processing as perceived by cross-checked observers of video-taped recordings of the lecture and a sample of visits. Listed contributions from practitioners for refining the guidelines. Questionnaire evidence of the overall knowledge of recipients of local infection guidelines. General practitioners' attitudes to guidelines and the management of common infections by questionnaire survey. The itemised prescribing of general practitioners from Prescribing Pricing Authority data. Special investigation activity of the practitioners from laboratory data.

Patient outcomes for a common infection by prospective clinical audit.

*Method*

A three-armed, minimised controlled trial. All practices in the district (bar one exclusion) were empanelled. The three groups of practices were selected by a process of correctional assignment (minimisation) to balance known major confounding factors. Of the three groups:

| **Group** | **Intervention** |
|---|---|
| Control | no intervention |
| Superficial | receipt of guidelines and supporting passive lecture event. |
| Deep | receipt of guidelines and engagement in deep processing by means of face to face interviews. |

Each intervention lasted 1 hour and the educational agenda and objectives were identical. Interventions were performed by two separate teams of two doctors; one team experienced in lecturing, the other familiar with deep processing theory. Superficial and deep processing interventions were applied synchronously and completed within 8 weeks. Data was analysed for the year prior and year subsequent to the start of intervention (1st November 1993).

This chapter describes how the methodology described in the previous chapter was applied in practice. Finally the strengths of the process and problems encountered are examined.

## 6.1 Purpose of the experiment.

To determine whether a deep processing approach to developing and implementing clinical guidelines is a) effective, and b) more effective than a standard superficial approach.

## 6.2 Setting

The Wirral peninsula, with a population of 350,000, is served by 180 general practitioners and one major acute hospital. The peninsula is bounded by the Irish Sea, River Dee, River Mersey and the Cheshire Plain. The main conurbation is Birkenhead which faces Liverpool to the North and is connected to that City by road and rail tunnels, and a passenger ferry. The Mersey side is heavily populated and industrialised. In contrast the Dee side is sparsely populated, the community composed of more affluent commuting professional people. There is little flow of population across the boundaries. The population is served by a major acute (NHS) trust hospital and there are few secondary services sought by Wirral residents elsewhere. The Wirral district is divided into 6 administrative areas. Social deprivation scores (from national census) and standardised mortality rates vary markedly from locality to locality. The standard of general practitioner service is generally high. However, there is a considerable range: the district contains both an Audit Commission prescribing excellence exemplar practice and a nationally recognised poor prescriber, the two ends of the spectrum.

## 6.3 Subjects

The subjects were all 69 general practices in Wirral district (i.e. currently on the list of Wirral Family Health Services Authority). There were a mean 2000 patients per practitioner and a mean 3 practitioners per practice. Every practice had practice nurse support. All microbiological specimens were processed by the microbiology laboratory at Wirral Hospital (NHS) Trust. Every NHS prescription dispensed was recorded on the National database of the Prescription Pricing Authority. Clinical audit was voluntary in general practice and supported by the Wirral Medical Audit Advisory Group (MAAG) funded by the Wirral Family Health Services Authority (FHSA). Hospital services were contracted for and funded by the Wirral District Health Authority (DHA). The FHSA and DHA were amalgamated since 1991 with a joint chief executive officer. A quarter of Wirral practices were 'fundholders' and contracted for non-emergency hospital services directly. They had budgets for prescribing cost, practice staff and hospital services; savings could be retained by practices to invest in premises improvements or new patient services. A third of practices were single-handed and tended to be more isolated and idiosyncratic in the way they practiced. Some practices, having achieved appropriate standards of organisation, were designated

training status and were provided with a supernumerary 'trainee' GP.

## 6.4 Selection and minimisation process

Each practice was assigned a code number and index card. The possible confounding variables were marked on each card:

- Single-handedness
- Fundholding status
- Training status
- Geographical location
- Overall prescribing characteristics

An independent research assistant mixed the anonymised cards thoroughly and assigned each to one of three experimental classes. The cards were assigned deliberately to groups so as to minimise* any imbalances in the proportions of variables across the groups.[188] Finally the grouped cards were placed in three plain unmarked envelopes and randomly allocated their final classification as the control, superficial, and deep groups (figure 6.1).

The final balance (after exclusion of the clinical tutor's practice; he was aware of the study objectives) was as follows:

*Minimisation seeks to distribute known confounding variables equally where stratification before randomisation would be impossibly complex due to multiple confounding modalities.

| Variable | Control group | Superficial group | Deep group |
|---|---|---|---|
| **Single-handed** | 6 | 5 | 6 |
| **Fundholding** | 7 | 7 | 6 |
| **Training practice** | 5 | 6 | 5 |
| **Locality A** | 3 | 2 | 4 |
| **Locality B** | 3 | 3 | 3 |
| **Locality C** | 4 | 4 | 3 |
| **Locality D** | 4 | 4 | 4 |
| **Locality E** | 3 | 4 | 3 |
| **Locality F** | 5 | 6 | 6 |
| **Low volume, high cost drugs** | 7 | 6 | 6 |
| **Low volume, high cost drugs** | 6 | 5 | 7 |
| **High volume, high cost drugs** | 3 | 5 | 3 |
| **High volume, low cost drugs** | 6 | 7 | 7 |
| **Number of practices** | 22 | 23 | 23 |

The randomisation and minimisation approach achieved as good a balance as stratification could have done without incurring unnecessary complexity. With so many known confounding variables some are bound to coexist in most practices, perfect matching was therefore impossible. However, with regard to prescribing, important variables seemed to be balanced satisfactorily across the groups. It was decided to exclude the clinical tutors practice from the study because of his knowledge of the research, hence one group (control) is smaller than the others.

During 1994 the learning styles of GPs in the district were ascertained through the Honey and Mumford Learning Styles

Questionnaire - a standard tool for studying adult education.[189] Their readiness to learn was simultaneously ascertained using the Self-Directed Readiness to Learn Scale (SDRLS), an American tool developed by Guglielmino.[190] The SDRLS establishes learning preferences solely from the subject's perspective; the Learning Styles Questionnaire examines educational preferences in terms of context and content - using both types of tool established a more complete understanding of the educational characteristics of the subjects.[191]

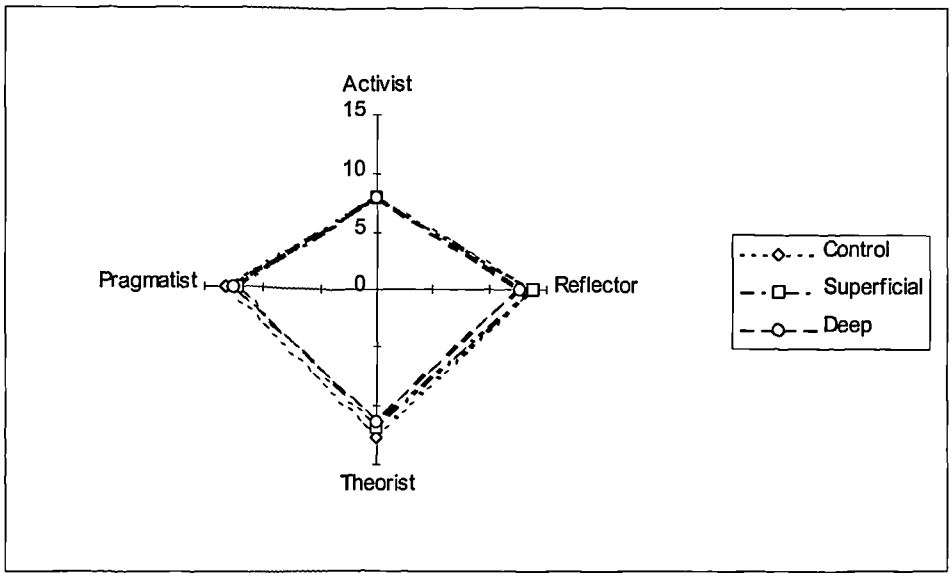



Figure 6.4a Preferred learning styles in the three groups; the patterns are similar.

The groups were educationally similar in terms of both preferred learning style* and readiness to learn† (figures 7.1a and 7.1b). Response rates to the two educational questionnaires was 63% (113/180) and 62% (112/180) respectively.

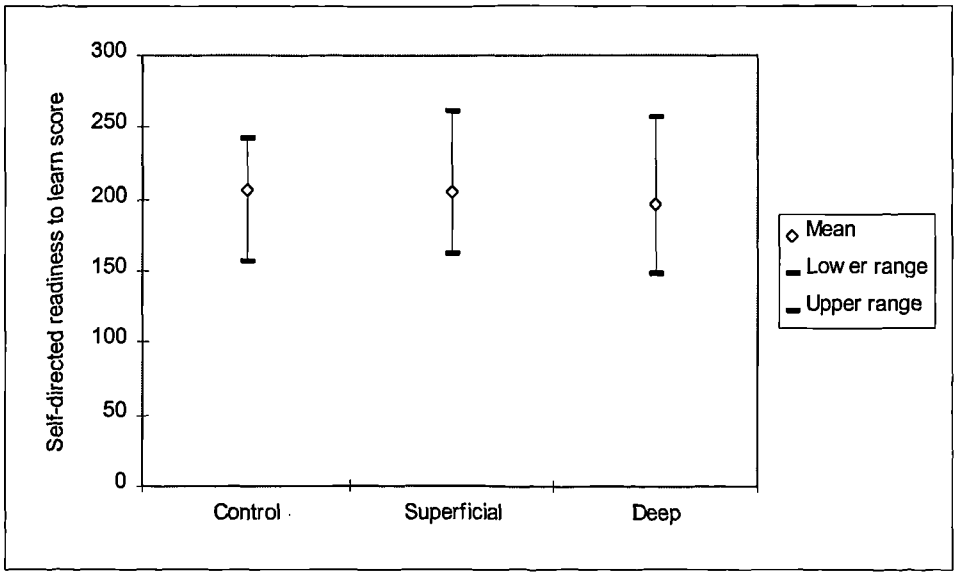



Figure 6.4b Readiness to learn scores in the three groups; mean self-directed readiness to learn and spread of scores is similar. The mean (not median) is shown as in the original studies in the USA by Gugliemino.

It may therefore be asserted that any differences in knowledge, attitudes, or behaviour after intervention cannot be

*Activists like novel experiences and are open-minded, reflectors like to ponder, theorists like logical complexity, pragmatists like to try things out; every learner is a mixture of these extremes (Honey and Mumford 1986).
†The self-directed learning readiness scale is a measure of how motivated individuals are to learn (Guglielmino 1977).

due to differences in learning style or readiness to learn between the three groups.

## 6.5 Study design

The effectiveness of the standard superficial approach was not known. The control group therefore provided a baseline to compare the relative effectiveness of the two interventions against. Hence the three arms to the trial. Essentially, it was a *three-armed, prospective, minimised, controlled trial.* The nature of the interventions rendered 'blinding' of the researchers and subjects impossible. This made adherence to this study protocol particularly important if observer bias was not to be introduced.

## 6.6 Ethical considerations

There were no direct interventions on patients and no-one was deprived of necessary treatment, therefore patient 'consent' was unnecessary. However, the doctors were not self-selected, some of them were denied access to guidelines against their will. The patients and doctors excluded from the deep group were being denied potentially better treatment and education. The proposed project was duly submitted to the Wirral Research Ethics Committee and was granted a Certificate of Approval on 23rd July 1993.

## 6.7 Materials

The essential requirement was for a set of scientifically derived guidelines on infections in the community. This was developed by Dr CA Barzokas - consultant medical microbiologist - and myself with a panel of expert advisers over a period of approximately two years.[192] For the purposes of the study The guidelines were published in large format to facilitate reading and provide space for recipients to write comments, notes and corrections.

A portable computer and appropriate software was employed to collect and transport data between the various organisations involved in the project.

## 6.8 Intervention

The **control group** were excluded from all activities associated with the study. They did not receive guidelines, invitations to lectures or offers of practice visits. They remained largely unaware of the study. The response to requests for guidance on infection management from this group was to direct them to other more generally available sources. Such requests were usually made directly to the local consultant medical microbiologist (Dr CA Bartzokas) or the local medical adviser (myself).

The **superficial group** received a copy of the clinical guidelines by postal delivery in the first week of the study year (1st November 1993 to 31st October 1994). An invitation to a lecture event to launch the guidelines was enclosed. The event was scheduled for 1.00pm to 2.00pm on Friday the 26th November and was preceded by a tempting light buffet for the attenders; it mimicked the usual pharmaceutical company sponsored educational events. Telephone calls to the practice managers were made to ensure that their practices were represented by at least one partner. Ultimately, 22 out of a possible 60 general practitioners (from the 23 practices) in the group attended. The hour was divided into a 25 minute lectures on community acquired

infections and their management. A senior lecturer in medical microbiology spoke on the microbiological aspects, and a senior lecturer in clinical pharmacology on the therapeutic aspects. Both lecturers were from the University of Liverpool and were experienced educators. They used the lectures to promote the guideline content and to encourage the audience to feedback comments and criticisms to the authors. The lectures were strictly in didactic style[193], but a question and answer session took place in the final 10 minutes of the event to allow a realistic amount of interaction. The event was designed and organised so as to mimic the typical current GP seminar or pharmaceutical company promotional event. Postgraduate education allowance approval was sought and achieved in the usual way. There was no further contact with this group on the matter for the rest of the intervention year.

The **deep group** received a copy of the guidelines by postal delivery at the same time as the superficial group. The deep group were not invited to the lecture event, but instead were informed that the local medical adviser and consultant medical microbiologist would be contacting them to arrange an appointment to discuss the guideline content and to receive their criticisms and comments. Every practice (23 in all) in the group

was visited in the first 12 weeks of the study year. The format was similar to the standard medical adviser approach with specific refinements. The first half of the visit was designed to *motivate* the practitioners to deeply process the guideline messages, the second half was intended to involving *deep processing* exclusively. A strict agenda based on two current and psychological theories (see below) was followed. Some superficial processing was utilised in the first half to facilitate the engagement of deep processing. The meeting which lasted for one hour closed with a verbal commitment by the practitioners to further scrutinise and try out the guideline content and to feedback the information in a subsequent meeting sometime in the future. Commitment was always achieved. Follow-up visits took place at the end of the intervention year (October and November 1994).

There was little evidence of contamination between groups; one GP in the deep group requested guidelines on behalf of her husband who was in the control group (we advised her that all Wirral GPs would receive a copy in due course) otherwise there was no evidence of sharing of the material.

## 6.9 Educational agenda - shared messages

It was essential that the educational objectives of both the lecturers and the practice visitors were identical even though the methods employed would be different. Otherwise like would not be compared with like. For example, if the visitors adhered to the guideline advice and the lecturers gave alternative advice it would be spurious to attribute subsequent differences in behaviour to different depths of processing. Similarly, identical arguments to support the guideline content was followed. For instance, the promotion of co-amoxiclav in pneumonia would have to be justified on the same grounds.

To achieve this both the lecturers and the visitors were contributors to the guidelines and therefore shared opinions and rehearsed the arguments together. The topics covered in the second half of the practice visits were the same as covered in the lectures, these were urinary tract infection, lower respiratory tract infection, pelvic infection, the appropriate use of antibiotics and laboratory investigations, and the avoidance of unnecessary new antibiotics. The educational objectives of both the lectures and visits were the same and were stated at the end of both types of intervention. These were to make practitioners aware of the guideline content, to encourage them to scrutinise and apply

them, and feedback their criticisms and comments to the guideline authors.

The lectures covered the topics above faithfully. Both the lecturers and the audience were videotaped simultaneously using a split-screen method to aid analysis. The lectures ran slightly over the time scheduled (60 minutes) because the question and answer session was not closed until every question which arose had been addressed by the lecturers.

The visits also covered the above topics, but a strict agenda was devised and followed. Peripheral cues, such as 'expert' status, were used to gain the attention of the practitioners. Alignment of the guidelines with the anticipated personal goals of the practitioners, such as mastery of their craft, was attempted. Emotional excitement was raised by reference to outside threats to the medical professional, such as greater Government control. Relevance of the issue to their professional lives was established, as was their capability to make competent and valuable criticisms and suggestions for the improvement of the guidelines. By now the practitioners were keen to deeply process the issues and open questions were raised by both parties. Finally, expressions of commitment to try out the guidelines and to feedback their

comments at a later visit were obtained from the practitioners. The actual agenda was as follows:

Table 6.9a The agenda for the deep processing visits.

| Agenda item | Aim: | Features: |
|---|---|---|
| 1. Introduction | Peripheral route | Peripheral cues:<br>• 'Expert' authors.<br>• Authority.<br>• Attractive book. |
| 2. Declaration | Alignment of goals (mastery of craft, self-perception, social responsibility, self-determination). | • Better medicine.<br>• Importance of GPs.<br>• Useful in practice.<br>• Puts GPs in control.<br>• Better for patients.<br>• No wastage of public funds. |
| 3. Importance | Emotional arousal | • Threat of imposition of case-management and policy-making from the USA.<br>• Managers concerned about the variability in clinical practice.<br>• Health economists now making the decisions instead of doctors.<br>• 'Expert' guidelines patronising to GPs. |
| 4. Relevance | Central route | • You are the practising doctors and need to consider these matters.<br>• Your input is as valuable as the research evidence.<br>• Other doctors will benefit from your contribution. |
| 5. Capability | Personal agency beliefs | • The book is explicit and easy to read.<br>• You deal with infections successfully all the time.<br>• Your comments will be taken seriously. |
| 6. Elaboration | Deep processing | • Any comments on clinical guidelines?<br>• Any comments on 'Microbes'?<br>• Discussion on specific infections. |
| 7. Close | Commitment | • Will you use 'Microbes'?<br>• Will you help us to improve it?<br>• Can we see you again in a few months to collect your thoughts? |

Items 1, 4, and 6 were drawn from the Elaboration Likelihood Model of persuasion of Richard Petty and John Cacioppo.*[194] This contends that once alerted to an issue by affective cues (superficial appearances etc.) if the issue seemed personally relevant and interesting then deep processing (effortful thinking on the issue) would follow resulting in attitudes associated with consistent behaviour. Items 2, 3, and 5 followed the doctrine of the Goals, Emotions, and Personal Capability Beliefs Theory of Motivation of Martin Ford.[195] He asserted that people could be motivated to do a thing they felt capable of if they could be excited enough and it would help them achieve a personal goal. The goals in item 2 were based upon the investigations of Colin Bradley into uncomfortable clinical situations in general practice, where it could be reasonably assumed that doctors would feel most uncomfortable where progress towards a personal goal had been frustrated.[196] Item 7 was promoted by the Audit Commission as good practice in practice visiting[197] and could be traced back to the work of social psychologists in the early 1970's.[198]

Agenda delivery was rehearsed with aid of video-recording and stop-watch. Both visitors were familiar with the agenda and delivered different parts of it. It was essential not to over-run the

---

* Petty and Cacioppo are American research psychologists in the field of social

time allocated to the lectures or any different impact on future clinical management in the group could be due to a greater quantity of educational input rather than a difference in the method. With so many practice visits to accomplish the intervention took most of the first 3-month period to achieve. A selection of four practices were video-taped in order to audit adherence to the agenda and measure the degree of deep processing achieved. The video-taped practices were selected according to availability of the recording crew and suitability of the premises for the television equipment. The selection was haphazard rather than strictly random, selection necessarily tended to be biased towards the larger practices with their larger premises (to accommodate the television equipment). A written log was kept of all visits.

## 6.10 Main measurables

The project was intended to improve the quality of clinical care in infection. This included the adoption of more effective and efficient treatment and investigation options. The rate of microbiological laboratory investigation was recorded for each practice. All the Wirral practices used Wirral Hospital (NHS) Trust laboratories exclusively so data capture was complete. The number of each type of investigation, its indication, and previous choice of antibiotic was collected automatically onto a database. The variables were linked to individual cases which allows direct conclusions rather than inferences to be drawn. However, the main weakness is the poor quality and incompleteness of the data on the forms; 30% are without clinical details. Sometimes even the name and address of the referring general practitioner are omitted. Even so, the large volume of investigations over the year might still enable useful comparisons to be made.

The number and type of prescriptions for antibiotics was routinely collected as PACT (prescribing activity and cost) data by the Prescription Pricing Authority. This data related to items . dispensed, not prescribed, and was not linked to individual diagnoses. Despite these shortcomings the data was collected with great accuracy and again the great volume would allow useful

comparisons to be made. From a behavioural point of view the product name ordered on each prescription was of primary interest. The number of individual prescriptions of specific drugs ('items') was the most appropriate measure, rather than the quantity of tablets or some other denominator such as 'defined daily doses'. Cross-sectional comparisons would correct for temporal drifts in the quantity per item. PACT assigned an indicative price to each prescription which would enable economic analysis.

Changes in clinical behaviour should have followed changes in attitude. Questionnaire surveys of general practitioner's attitudes regarding choice of empirical antibiotic for common infections were distributed before and after the study year.

In order to examine in more detail the impact on individual cases an audit of management and outcome of uncomplicated cystitis in women was performed on (self-selecting) samples from each group. Because individual patient data was involved, the audit was performed in co-operation with the Wirral MAAG (Medical Audit Advisory Group - the local general practice audit organisation) which is registered under the Data Protection Act. The practitioners were self-selecting, but equally self-selecting from each group, and were blinded insofar as they were engaged

in a clinical audit and were unaware the results would also be analysed as the three experimental groups.

## 6.11 Timescale

The pre-study period of draft guideline development and planning took two years to complete. The interventions and the observation year was a full 12 months (divided for data analysis purposes into 3-monthly periods). Data analysis and evaluation took over 6 months to complete the revised version of the guidelines would take at least a further 6 months to achieve completion. The entire project was therefore expected to have taken at least 4 years to complete.

## 6.12 Data management

The prescribing data alone occupied 20 megabytes of computer disc space (the equivalent of over 10 paperback books). The laboratory data was of similar proportions. Even using iterating programs (called 'macros' in Microsoft Excel spreadsheets) a simple search took hours to complete. For example, the simple condensing of the data into sub-totals according to therapeutic class took over 6 hours to complete with a macro-program. Records were also kept of every visit on a written log and on video-tape in order to confirm consistency of approach and to estimate the depth of processing achieved. The questionnaire forms completed by the general practitioners on antibiotic prescribing attitudes and learning styles also had to be collated and stored in a retrievable way.

## 6.13 Validation of the process

Independent experts (research psychologists) were engaged to evaluate how well the protocol had been followed. The video-recordings of the visits were examined by 2 psychology postgraduate researchers to observe whether a) the agenda was adhered to, b) deep processing was achieved, and c) which parts of the approach worked badly or well. The eclectic nature of the project left it vulnerable if the design reflected the requirements of one discipline (i.e. general practice research).

The two researchers found that the agenda was adhered to almost too well in the visits, the lengthy introduction took up a good deal of the hour and reduced the scope for contributions from the GPs.

Body language interpretation was impossible as the view of the GPs' was often obscured and frequently the cameraman would zoom in on a speaker so the movements and posture of the others was hidden. However, an attempt was made to evaluate this aspect; inter-observer agreement was 76% (positive attention skills) and 68% (negative), both too low to be reliable.

The verbal evidence of deep processing was more reliable, however (see table 6.13a).

| Statements | Correlation (r) |
|---|---|
| New / novel | 0.9956 |
| Rephrase | 0.984 |
| Repeated | 0.4072 |
| Other relevant | 0.8931 |
| Other irrelevant | 0.9988 |

Table 6.13a Correlation coefficients (r) for the two scores of the 2 observers for each of the verbal indicators of deep processing.

Between the 2 observers, correlation coefficients (r) were very highly statistically significant ($P < 0.0005$) for every indicator except 'repeated statements' ($P > 0.05$).

## 6.14 Adherence to protocol.

The protocol was strictly adhered to. There was no examination of any individual practice's data before the end of the study year. The follow-up visits were delayed until the end of the year so as not to introduce the effect of repetition or feedback. Distribution of the guidelines was restricted to one per named participant doctor. The cystitis study was commenced after the main study year so as not to confound the experiment.

# Chapter 7

# Findings

*The totality of our so called knowledge...from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only along the edges. Or, to change the figure, total science is like a field of force whose boundary conditions are experience.*[199]

Willard Quine. 1951

## Summary

The deep group showed verbal evidence of deep processing 5 times longer (237 seconds, 6.5% of the time available) per GP than the superficial group (45 seconds per GP, or 1.2% of the time available), and showed a deeper (novel statements) quality of processing than the superficial group (relevant statements).

Knowledge of the guidelines improved by 21% in the deep group over the 12 month post-intervention study period after deep processing (Median difference in score 2/12; P=0.002, 95% CL = 1 to 3). Superficial processing had no effect (Median difference = 0/12; P = 0.6 95% CL = -1 to 1).

Attitudes regarding treatment of each of four common infections was more favourable than control following deep processing and all changes in attitude following deep processing were significant, whereas there were no significant changes following superficial processing.

Summary table 1. Changes in attitude towards the treatment of some common infections following intervention (control = non-intervention). Negative result (-) indicates an adverse change.

| Infection | Control group | Superficial group | Deep group |
|---|---|---|---|
| **Cystitis** | 10% (P = 0.3) | 12% (P = 0.22) | 28% (P = 0.001) |
| **Acute Otitis Media** | -4% (P = 0.83) | 4% (P = 0.83) | 26% (P = 0.02) |
| **Bronchopneumonia** | 14% (P = 0.3) | 5% (P = 0.85) | 38% (P = 0.001) |
| **Sinusitis** | -10% (P = 0.45) | 13% (P = 0.22) | 25% (P = 0.02) |

For relevant antibiotics their share of total prescriptions was more favourable in the deep group than control; advocated antibiotics were prescribed significantly more frequently in the deep group than the superficial following intervention (except in the case of co-amoxiclav), and antibiotics advised against (co-trimoxazole, tetracyclines, newer antibiotics) were prescribed less frequently in deep than control.

Summary table 2. Changes in the proportion of several antibiotics (as proportions of total antibiotic prescriptions) after intervention.

| Infection | Control group | Superficial group | Deep group |
|---|---|---|---|
| Amoxycillin | 11% (P <0.0001) | -3% (P = 0.0002) | 6% (P <0.0001) |
| Co-amoxiclav | 29% (P <0.0001) | 38% (P <0.0001) | 32% (P <0.0001) |
| Trimethoprim | 4.2% (P = 0.002) | 7.4% (P <0.0001) | 14.5% (P <0.0001) |
| Newer antibiotics | 19% (P <0.0001) | 14% (P <0.0001) | -5% (P = 0.04) |
| Co-trimoxazole | -27% (P <0.0001) | -22% (P <0.0001) | -32% (P <0.0001) |
| Tetracyclines | -3% (P = 0.1) | 1% (P = 0.7) | -8% (P <0.0001) |

Summary table 3. Differences (proportions of total antibiotic prescriptions) between the groups after the intervention.

| | Superficial v. Control | Deep v. Control | Deep v. Superficial |
|---|---|---|---|
| Amoxycillin | 0.3% (P = 0.15) | 1.7% (P < 0.0001) | 1.4% (P < 0.0001) |
| Co-amoxiclav | 0.6% (P <0.0001) | 0.1% (P = 0.38) | -0.5% (P <0.0001) |
| Trimethoprim | 0.8% (P < 0.0001) | 1.4% (P < 0.0001) | 0.6% (P < 0.0001) |
| Newer antibiotics | -0.4% (P < 0.0001) | -0.4% (P <0.0001) | -0.1% (P = 0.04) |
| Co-trimoxazole | 2.8% (P < 0.0001) | 0.9% (P = 0.16) | -1.9% (P = 0.002) |
| Tetracyclines | 0.19% (P = 0.047) | -0.7% (P < 0.0001) | -0.89% (P < 0.0001) |

The guidelines advocated more mid-stream specimen of urine, less gynaecological, and more gastro-intestinal microbiological investigations.

The superficial group performed significantly less mid-stream specimen of urine investigations after intervention than control (-1.5%, P = 0.02); the deep was similar to control. Gynaecological investigations were performed significantly less in the deep group than in either control or superficial following intervention (-3.7%, P < 0.0001; and -3.99, P = < 0.0001 respectively). The deep group performed more intestinal investigations than either control or superficial following intervention (1.19%, P = 0.0004; and 1.65%, P < 0.000 respectively).

The increase in cost per antibiotic prescription after deep processing was 3p less than the control group (1.6% of cost per prescription in 1993), a saving of £9,160 on antibiotic costs in that group, whereas increase following superficial processing was 5p (2%) more.

In a clinical study of cystitis 12 months after intervention the relapse rate after initial consultation was significantly less likely for patients of deep group doctors than either control or superficial (8.9%, P = 0.03; and 14.6%, P = 0.006 respectively).

## 7.1 Measurement of deep processing

### Deep processing in the deep processing group

*Level of contribution:*

Verbal contributions were always audible. They were examined for quantity (time elapsed) and quality (type of contribution). Content of contributions to discussion can be ranked according to the level of processing they indicate,* from deep processing to superficial they are:

1. new or novel
2. rephrase
3. repetition
4. other relevant
5. other irrelevant

*Deep*

*Superficial*

Inter-observer agreement was satisfactory with high correlation for all categories (Pearson's product moment correlation $P < 0.0005$, $r = +0.8931$ to $+0.9988$) except 'repeated statements' ($P > 0.05$, $r = +0.4072$).

* I am grateful to Prof. PD Slade and colleagues of the University of Liverpool for allowing the use of this research tool, currently unpublished.

The mean readings of both observers are shown in table 7.1a. Repeat statements took up only a few seconds and so the poor inter-observer correlation on that variable does not significantly affect the instrument adversely.

Table 7.1a The time (minutes or seconds) each 'deep processing group' practice spent in discussion.

| Practice | New | Re-phrase | Repeat | Relevant other | Irrelevant other | Total | Per GP (mean) |
|---|---|---|---|---|---|---|---|
| 1 | 5 mins. | 43 secs. | 6 secs. | 4 secs. | 0 | 6 mins | 2 mins. |
| 2 | 13 mins. | 2 mins. | 0 | 0 | 0 | 15 mins. | 5 mins. |
| 3 | 11 mins. | 4 mins. | 6 secs. | 3 mins. | 1 min. | 19 mins. | 6 mins. |
| 4 | 9 mins. | 1 min. | 5 secs. | 5 secs. | 2 mins. | 12 mins. | 3 mins. |

It is evident that the contributions were mainly at the deep processing end of the continuum and that the contribution time per GP exceeded that in the superficial (lecture) group (mean 45 seconds per GP).

Table 7.1b Verbal evidence of deep processing during the two interventions.

| **Experimental group** | **Usual contribution type** | **Time contributing per GP** |
|---|---|---|
| **Superficial:** | Relevant other | 45 seconds |
| **Deep:** | Novel / new comment | 237 seconds |

It is apparent that significantly more and significantly deeper processing was achieved in the deep processing group.

## 7.2 Suggestions for improvement of the guidelines

The suggestions and criticisms received from the deep group are further evidence of deep processing in that group. No suggestions were received from the superficial group. Suggestions for improvement in the deep group arose at first and second interview equally. A total of 52 suggestions were received.

Table 7.2a Summary suggestions for improvement received from the general practitioners in the deep group (overleaf).

| Topic | Suggestion | No. |
|---|---|---|
| Appendices | Advice on overnight storage of specimens. | 1 |
| Central nervous system | Options for pre-admission treatment of meningitis. | 1 |
| Cardiovascular system | Add advice on dental prophylaxis etc. | 2 |
| Dental & oral | Should dental abscesses be treated with co-amoxiclav rather than Amoxycillin (same bacteria as LRTI).<br>Add advice on parotitis. | 1<br>2 |
| ENT | Refer suppurative otitis media only if persistent.<br>Otitis externa - when to use drops and when to refer.<br>Sinusitis - does doxycycline penetrate better than co-amoxiclav? | 1<br>1<br>1 |
| Eyes | Advice on sticky eyes, e.g. swab and massage of duct.<br>No need to swab every case of conjunctivitis, only if question of STD. | 1<br>1 |
| Format | Thumb indents or similar.<br>Separate summary sheet(s).<br>Hand-held size, e.g. A5. | 1<br>1<br>2 |
| General | Each infection should be defined.<br>Add index.<br>Add section on leptospirosis.<br>Add section on parasites and fungi<br>Add section on viruses and exanthemata.<br>Ensure principal investigations are applicable to general practice. | 1<br>2<br>1<br>2<br>2<br>1 |
| Gastro-intestinal | Add section on helicobacter pylori.<br>Advice on giardia.<br>Step by step approach to diarrhoea. | 5<br>1<br>1 |
| Guidance notes | Infections in pregnancy - significance, investigations and treatment.<br>Infections in childhood - appropriate treatment and doses. | 1<br>1 |
| Genito-urinary | Name of good bladder irrigation fluid.<br>Clarify epididymo-orchitis, perhaps with PID.<br>Advice on recurrent sterile pyuria. | 1<br>1<br>1 |
| Gynaecological | Practical advice on PID / vaginitis.<br>No need to investigate every time, treat PID with potent antibiotic<br>e.g. oxytetracycline or co-amoxiclav with metronidazole and refer if it recurs.<br>Bartholinitis - treat with flucloxacillin. | 6<br>1<br><br><br>1 |
| Respiratory | Criteria for pneumococcal, influenzal vaccination.<br>Reasons for referring pleurisy.<br>More options for therapy e.g. 2nd and 3rd line.<br>Pre-admission treatment for pneumonia only necessary if delay in admission expected e.g. rural practices.<br>Classify chest infections; age, COAD etc. | 1<br>1<br>1<br>1<br><br>1 |

Table 7.2b Summary suggestions for improvement received from the practice nurses in the deep group.

| |
|---|
| Hand cleansing and infection control. |
| Sterilisation of skin before injections. |
| Expelling air before injection. |
| Clearer separation of management and procedures. |
| Storage of specimens. |
| Refrigeration advice. |
| Storage of specimens. |
| Control and storage of hazardous substances. |
| Hepatitis B vaccination. |
| Protection against HIV and AIDS. |
| Polio vaccination. |
| Vaccination advice. |

The above tables strictly exclude comments arising purely during discussion. Usually practitioners had made a note of their comment prior to interview. It was rare for GPs to comment on procedures, or for practice nurses to comment on clinical management. The greatest conflict between science and practice is clearly in the management of pelvic infection where the 'investigate before treatment' recommendation in 'Microbes' was felt to be untenable in mild cases in general practice. Particularly it was felt inappropriate to leave a Friday evening patient suffering over a weekend whilst awaiting investigation.

## 7.3 Impact on knowledge.

The questionnaire at the end of the study year asked practitioners their choice of antibiotic before investigation. The guideline recommendations can be adopted as the gold standard against which to measure the performance of each respondent. 'Perfect' responses would therefore be:

Table 7.3a Guideline recommendations (and hence 'correct' answers) for the selection of infections inquired about.

| Question | Infection | Blind treatment (answer) |
|---|---|---|
| 1 | Upper respiratory tract infection | None |
| 2 | Tonsillitis | Penicillin V (or erythromycin) |
| 3 | Bronchitis | Co-amoxiclav (or erythromycin) |
| 4 | Bronchopneumonia | Co-amoxiclav (or erythromycin) |
| 5 | Lobar pneumonia | Co-amoxiclav (or erythromycin) |
| 6 | Acute otitis media | Amoxycillin (or erythromycin) |
| 7 | Cystitis | Trimethoprim (or cephalexin) |
| 8 | Sinusitis | Co-amoxiclav (or erythromycin) |
| 9 | Impetigo | Mupirocin +/- Flucloxacillin or erythromycin. |
| 10 | Cellulitis | Flucloxacillin or erythromycin. |
| 11 | Gastro-enteritis | None |
| 12 | Salpingo-oophoritis | Refer for full investigation |

Maximum score was 12/12. No-one scored 0 or 12 although every question was answered correctly by at least one respondent. The actual performance of each group was:

Table 7.3b The median scores of each group (range = 0 to 12).

| | Control | | Superficial | | Deep | |
|---|---|---|---|---|---|---|
| | Before | After | Before | After | Before | After |
| Median score | 7 | 7 | 7 | 7 | 7 | 8.5 |

Table 7.3c Differences in scores before and after intervention. Analysis is by Mann-Whitney test and confidence limits are for the difference between medians.

| | **Median difference** | **P** | **Lower 95%CL** | **Upper 95%CL** |
|---|---|---|---|---|
| ***Within group differences before and after interventions:*** | | | | |
| Control | 0 | 0.91 | -1 | 1 |
| Superficial | 0 | 0.55 | -1 | 1 |
| Deep | 2 | 0.002 | 1 | 3 |
| ***Between groups before interventions:*** | | | | |
| C *versus* S | 0 | 0.84 | -1 | 1 |
| C *versus* D | 0 | 0.4 | -1 | 1 |
| S *versus* D | 0 | 0.5 | -1 | 1 |
| ***Between groups after interventions:*** | | | | |
| C *versus* S | 0 | 0.8 | -1 | 1 |
| C *versus* D | -1 | 0.007 | -2 | 0 |
| S *versus* D | 1 | 0.03 | 0 | 2 |

The groups are not significantly different prior to intervention. However, afterwards the deep group has increased its score significantly, the other groups show no significant change.

The distributions of the scores:

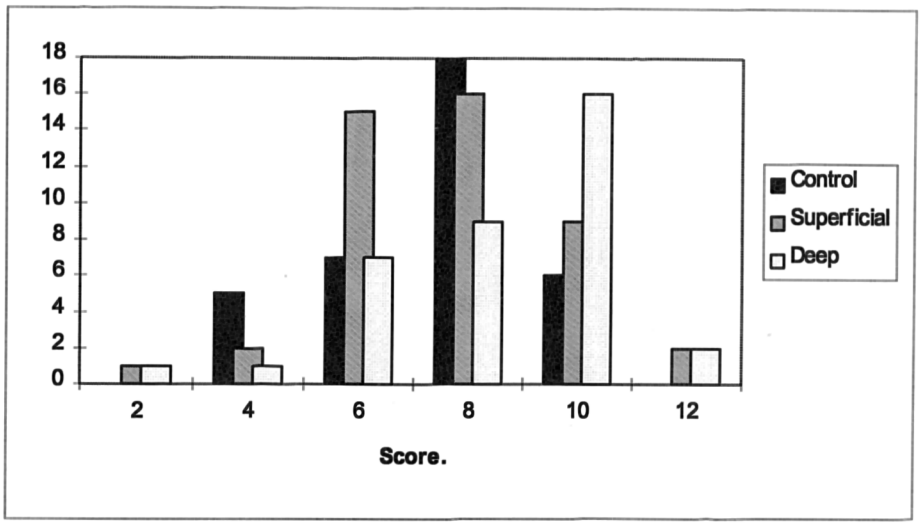



Figure 7.3a - The distributions of scores in the three groups. Scores are grouped in twos (i.e. 1 and 2, 3 and 4, etc.).

## 7.4 Differences in Attitudes.

Questionnaires on common infections were sent to all Wirral GPs prior to the interventions (Autumn 1993) and after the study year (Winter 1994). For doctors, response rates to both were high.[200]

Table 7.4a Response rates for the attitude questionnaires before and after the experiment. All 191 Wirral general practitioners were sent both questionnaires.

| | **Control:** | | **Superficial:** | | **Deep:** | |
|---|---|---|---|---|---|---|
| | before | after | before | after | before | after |
| Responses | 48 | 36 | 51 | 46 | 45 | 36 |
| Total in group | 61 | 61 | 64 | 64 | 66 | 66 |
| Response rate (%) | 79% | 59% | 80% | 72% | 68% | 56% |

However, despite a shorter questionnaire and reminder the response rates were not so high to the second questionnaire. Questionnaire 'fatigue' may have set in as GPs were being bombarded with marketing surveys from the developing NHS purchasers and trusts. For example one respondent remarked, "More paper. More time".

Before and after responses are not paired and the groups are not matched. Fisher's exact test is employed in this section as numbers are relatively small. Confidence intervals for odds ratios

are at 95% (Gart's method*)[201] as is the probability of the null hypothesis being true (P). Two-tailed tests are employed as a routine. Differences are expressed as percentages to ease appraisal, but the reader is asked to bear in mind that differences may appear exaggerated because n<100.

*A method of constructing exact confidence limits for 2x2 tables.

### 7.4.1 Cystitis treatment.

The GPs were asked their empirical (blind) antibiotic of first choice in cystitis. All groups increasingly adopted the recommended drug trimethoprim. Post-intervention, the difference between the superficial and control groups is negligible, but the deep group is significantly different. No significant difference is evident between the proportion of trimethoprim responses in the control and superficial groups either before or after. The post-intervention responses show a reduced range of antibiotics for cystitis in the deep group.

The guidelines recommended 3 day courses of antibiotics for uncomplicated cystitis, as had local mid-stream specimen of urine reports to GPs for several years. 50% of responses in each group (+/- 1%) were for 3 day courses. This question was only presented in the second questionnaire, so before and after effects are unknown. The control mean length of course was 4.39 days (95%CL = 3.84 to 4.93), superficial 4.13 days (CL = 3.72 to 4.55), deep 4.06 days (CL = 3.6 to 4.51), but the differences are not significant

Table 7.4.1a GPs expressing preference for trimethoprim as the empirical antibiotic in cystitis.

| | **Control:** | | **Superficial:** | | **Deep:** | |
|---|---|---|---|---|---|---|
| Antibiotic | before | after | before | after | before | after |
| Trimetho-prim | 73% (35/48) | 83% (30/36) | 73% (37/51) | 85% (39/46) | 69% (31/45) | 97% (35/36) |

Table 7.4.1b The difference between before and after intervention in each group.

| Within group comparison before *versus* after: | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | 10% | 0.3015 | 1.8572 | 0.5672 | 6.6802 |
| Superficial | 12% | 0.2167 | 2.1081 | 0.6962 | 6.8497 |
| Deep | 28% | 0.0011 | 15.8065 | 2.1122 | 687.0277 |

Table 7.4.1c The differences between groups before intervention.

| Between group comparison before: | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 0% | 1 | 1.0187 | 0.3829 | 2.7229 |
| C *versus* D | -4% | 0.82 | 1.2159 | 0.4501 | 3.2889 |
| S *versus* D | -4% | 0.8225 | 0.8225 | 2.2239 | 0.3162 |

Table 7.4.1d The differences between groups after intervention.

| Between group comparison after: | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 2% | 1 | 1.1143 | 0.2775 | 4.328 |
| C *versus* D | 14% | 0.1065 | 7 | 0.7639 | 330.9377 |
| S *versus* D | 12% | 0.073 | 6.2821 | 0.7334 | 291.06 |

Table 7.4.1e Preferred empirical antibiotic in cystitis: actual responses at the end of the study.

| Cystitis | Control | Superficial | Deep |
|---|---|---|---|
| Trimethoprim | 30 | 39 | 35 |
| Co-amoxiclav | 2 | 1 | |
| Cephalexin | 2 | 4 | |
| Co-trimoxazole | 1 | | |
| Cephradine | 1 | | |
| None | | 1 | |
| Macrodantin | | | 1 |
| Total (n) | 36 | 45 | 36 |

### 7.4.2 Acute otitis media.

Wirral GPs were asked which antibiotic was their (blind) first choice in acute otitis media. The guidelines promoted amoxycillin. The amoxycillin responses from the control group have reduced, those in the superficial have increased, and the deep group have increased. After intervention superficial positive response is higher than control, and deep higher than control and superficial, but not quite to the required significance level. Before and after differences are only significant in the deep group.

Table 7.4.2a GPs expressing preference for amoxycillin as empirical antibiotic in acute otitis media

| | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| Antibiotic | before | after | before | after | before | after |
| Amoxycillin | 54% (26/48) | 50% (18/36) | 55% (28/51) | 59% (27/46) | 47% (21/45) | 72% (26/36) |

. Table 7.4.2b The difference between before and after intervention in each group.

| Within group comparison before *versus* after: | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | -4% | 0.8258 | 1.1818 | 0.4554 | 3.0671 |
| Superficial | 4% | 0.8377 | 1.1673 | 0.4838 | 2.82501 |
| Deep | 26% | 0.0249 | 2.9714 | 1.0662 | 8.5175 |

Table 7.4.2c The differences between groups before intervention.

| Between group comparison before: | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 1% | 1 | 1.0301 | 0.4333 | 2.4479 |
| C *versus* D | -8% | 0.5362 | 1.3506 | 0.5529 | 3.306 |
| S *versus* D | -11% | 0.5397 | 1.3913 | 0.5771 | 3.3607 |

Table 7.4.2d The differences between groups after intervention.

| Between group comparison after: | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 9% | 0.5051 | 1.4211 | 0.5397 | 3.745 |
| C *versus* D | 22% | 0.0898 | 2.6 | 0.882 | 7.8254 |
| S *versus* D | 13% | 0.2989 | 1.8296 | 0.6555 | 5.2625 |

The actual post-intervention responses were as follows:

Table 7.4.2e Preferred empirical antibiotic in acute otitis media: actual responses at the end of the study.

| **Acute O.M.** | **Control** | **Superficial** | **Deep** |
|---|---|---|---|
| Amoxycillin | 18 | 27 | 26 |
| Co-amoxiclav | 9 | 13 | 6 |
| Ampicillin | 4 | | |
| Erythromycin | 2 | | |
| Penicillin V | 2 | | |
| Trimethoprim | 1 | 3 | 1 |
| Ciprofloxacin | | 1 | |
| Co-trimoxazole | | 1 | |
| Clarithromycin | | | 1 |
| Cefaclor | | | 1 |
| *Total (n)* | *36* | *45* | *36* |

### 7.4.3 Bronchopneumonia.

GPs were asked which was their antibiotic of first choice in bronchopneumonia. The guidelines promoted co-amoxiclav. The deep group practitioners showed less inclination than either of the other two groups to choose co-amoxiclav prior to intervention. Over the study period control, superficial and deep responses became more positive to co-amoxiclav; only the deep change was significant. At the end of the study deep responses were more positive than control or superficial. There was no significant change in either control or superficial. However, the deep group change was statistically highly significant. Neither intervention group was significantly different to the control post-intervention.

Table 7.4.3a GPs expressing preference for co-amoxiclav as empirical antibiotic in bronchopneumonia.

| | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| Antibiotic | before | after | before | after | before | after |
| Co-amoxiclav (total) | 50% (48) | 64% (48) | 56% (51) | 61% (46) | 29% (45) | 67% (36) |

Table 7.4.3b The difference between before and after intervention in each group.

| Within group comparison before *versus* after: | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | 14% | 0.2951 | 1.7692 | 0.6704 | 4.733 |
| Superficial | 5% | 0.847 | 1.1801 | 0.486 | 2.8767 |
| Deep | 38% | 0.0014 | 4.9231 | 1.7376 | 14.1696 |

Table 7.4.3c The differences between groups before intervention.

| Between group comparison before: | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 6% | 0.6295 | 1.3182 | 0.5548 | 3.1365 |
| C *versus* D | -21% | 0.0611 | 2.4615 | 0.964 | 6.3818 |
| S *versus* D | -27% | 0.0102 | 3.2448 | 1.2835 | 8.3355 |

Table 7.4.3d The differences between groups after intervention.

| Between group comparison after: | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | -3% | 0.9617 | 1.1374 | 0.4216 | 3.1047 |
| C *versus* D | 3% | 1.0 | 1.1304 | 0.3836 | 3.3426 |
| S *versus* D | 6% | 0.7588 | 1.2857 | 0.4719 | 3.5607 |

The actual post-intervention responses were as follows.

Table 7.4.3e Preferred empirical antibiotic in bronchopneumonia: actual responses at the end of the study.

| Bronchopneumonia. | Control | Superficial | Deep |
|---|---|---|---|
| Co-amoxiclav | 23 | 28 | 24 |
| Amoxycillin | 4 | 7 | 3 |
| Penicillin | 2 | | |
| Cephradine | 2 | | |
| Cefaclor | 1 | | 1 |
| Cefixime | 1 | | |
| Ciprofloxacin | 1 | 3 | 2 |
| Cephalexin | | 3 | 1 |
| Erythromycin | | 2 | 1 |
| Doxycycline | | 2 | |
| Not specified | 2 | | 1 |
| *Total (n)* | *36* | *45* | *36* |

### 7.4.4 Sinusitis.

The GPs' antibiotic of first choice in sinusitis. The guidelines recommended co-amoxiclav. Prior to intervention the deep group responses were significantly less positive than both control and superficial. Control positive responses fell over the study period, superficial and deep rose; only the deep group rose significantly. Ultimately both intervention groups' positive responses exceeded the control, but not significantly. The differences apparent before intervention had reduced. The actual post-intervention results show a narrower range of antibiotics in the deep group.

Table 7.4.4a GPs expressing preference for co-amoxiclav as empirical antibiotic in sinusitis.

| | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| Antibiotic | before | after | before | after | before | after |
| Co-amoxiclav | 23% | 13% | 20% | 33% | 6% | 31% |
| *Total* | *48* | *36* | *51* | *46* | *45* | *36* |

Table 7.4.4b The difference between before and after intervention in each group.

| Within group comparison before *versus* after: | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | -10% | 0.45 | 1.8432 | 0.5165 | 7.4723 |
| Superficial | 13% | 0.22 | 1.9839 | 0.7156 | 5.6325 |
| Deep | 25% | 0.02 | 3.4091 | 1.2329 | 9.6154 |

Table 7.4.4c The differences between groups before intervention.

| Between group comparison before: | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | -3% | 0.88 | 1.2189 | 0.4149 | 3.6063 |
| C *versus* D | -17% | 0.0005 | 5.0455 | 1.8876 | 13.7815 |
| S *versus* D | -14% | 0.0001 | 6.15 | 2.2692 | 17.1386 |

Table 7.4.4d The differences between groups after intervention.

| Between group comparison after: | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 20% | 0.09 | 3 | 0.8832 | 11.7306 |
| C *versus* D | 18% | 0.16 | 2.728 | 0.7411 | 11.2407 |
| S *versus* D | -2% | 1 | 1.0997 | 0.39 | 3.1586 |

Table 7.4.4e Preferred empirical antibiotic in sinusitis: actual responses at the end of the study.

| Sinusitis. | Control | Superficial | Deep |
|---|---|---|---|
| Doxycycline | 18 | 17 | 14 |
| Amoxycillin | 8 | 5 | 7 |
| Co-amoxiclav | 5 | 15 | 11 |
| Cephradine | 1 | | |
| Erythromycin | 1 | 2 | 1 |
| Oxytetracycline | 1 | | |
| Trimethoprim | 1 | 1 | 1 |
| Ampicillin | 1 | | |
| Tetracycline | | 1 | |
| Cephalexin | | 1 | 1 |
| Chlortetracycline | | 2 | |
| Co-trimoxazole | | 1 | |
| None specified | | | 1 |
| Total (n) | 36 | 45 | 36 |

**Attitude summary.**

The pattern emerges of a universal significant change in attitude in the deep group regarding common infections. The deep group was the only group to show significant changes in attitude and was more likely to choose the recommended drugs after intervention. Deep processing was associated with high positive attitudes towards the guideline drugs at the end of the study, and the most dramatic positive changes within the study year.

Table 7.4z Summary statistics on changes in attitude after interventions relative to control group changes.

| Group | Mean | 95% CL | Median | Minimum to maximum |
|---|---|---|---|---|
| Superficial | 6% | -15.2 to 27 | 5% | -9 to 23 |
| Deep | 26.8% | -15 to38.5 | 27% | 18 to 35 |

The subset of the superficial group who had attended both meetings and completed both questionnaires was similarly analysed (table 7.4zz):

Table 7.4zz Superficial group in total and the superficial group subset who attended the lectures and completed both questionnaires: Summary statistics for on changes in attitude after interventions relative to control group changes.

| Group | Mean | 95% CL | Median | Minimum to maximum |
|---|---|---|---|---|
| Superficial | 6% | -15.2 to 27 | 5% | -9 to 23 |
| Subset of attenders | 0.9% | -0.5 to 2.3 | 0.5% | -1 to 4 |

The changes in attitudes in the subset were no more favourable than the superficial group as a whole. This indicates that lecture attendance was not a predictive variable for favourable attitude changes and the low attendance at the lectures is not an explanation for the differences in attitude change in the deep and superficial groups.

## 7.5 Differences in prescribing.

Prescribing data is from PACT. It is accurate to within 0.1% with regard to the data transferred from the prescription forms to the data base. However it has some conspicuous weaknesses:

1. The population base upon which it makes calculations of prescribing rate is reliant upon the information from GP registers at the FHSA and is therefore suspect. For example in one quarter on Wirral a net 1500 people vanished - an unlikely event.
2. It measures only prescriptions dispensed.
3. None of the prescriptions can be linked to particular indications, cases, or even types of cases.
4. It is only meaningful down to practice-level because of prescription pad sharing within practices.

The more reliable denominator for prescribing data analysis is therefore total (or sub-total) items and not the registered population.

Chi squared tests are employed in this section in view of the large numbers involved; rather than exact tests which require tremendous computing power to cope with large factorials. Probability values are quoted after Yates' correction*,[202] 95% confidence intervals have been arrived at by the Cornfield

method[†203] and Haldane's correction[‡] applied. Firstly, the proportion of various guideline antibiotics (and some non-recommended antibiotics) as proportions of total antibiotic prescriptions are examined to determine whether significant changes in likelihood to prescribe them are apparent. Secondly, antibiotic prescribing by therapeutic class is examined for the final quarter of the study year to determine if there was persistence of any effects. Finally, changes in cost and rate of antibiotic prescribing per head of population are described. For interest and as a further control England data is included where available for the calendar years 1993 and 1994, and is thus only one month out of phase with the years before and after intervention.[204]

---

*A continuity correction (for greater precision) achieved by moving each actual number in a contingency table 0.5 nearer to its corresponding expected number.
†A reliable method for constructing confidence intervals for Chi squared tests.
‡0.5 is added to the value of each cell in a 2x2 table if one cell would cause a division by zero error.

### 7.5.1 Amoxycillin.

The guidelines recommended amoxycillin for acute otitis media and dental abscess. The superficial and deep groups were already significantly more likely to prescribe amoxycillin than control prior to intervention. The rate of prescribing of amoxycillin post-intervention is highest in the deep group. Post-intervention the likelihood of prescribing Amoxycillin is again significantly higher in deep than control. Further, the significant difference between superficial and deep evident after intervention was not present before.

Table 7.5.1a Amoxycillin prescriptions as proportions of total prescriptions in the year before and year after intervention.

| Annual figures | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| Antibiotic. | before | after | before | after | before | after |
| Total items | 125897 | 114549 | 134920 | 125479 | 145369 | 134884 |
| Amoxycillin | 16.3% | 18.1% | 18.9% | 18.3% | 18.7% | 19.8% |

Table 7.5.1b The difference between before and after intervention in each group: amoxycillin prescriptions.

| Within group comparison before *versus* After (Chi squared): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | 1.8% | <0.0001 | 1.132 | 1.108 | 1.156 |
| Superficial | -0.6% | 0.0002 | 1.038 | 1.018 | 1.059 |
| Deep | 1.1% | <0.0001 | 1.071 | 1.051 | 1.092 |

Table 7.5.1c The differences between groups before intervention: amoxycillin prescriptions.

| Between group comparison before (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 2.6% | <0.0001 | 1.193 | 1.169 | 1.217 |
| C *versus* D | 2.4% | <0.0001 | 1.178 | 1.154 | 1.201 |
| S *versus* D | -0.2% | 0.18 | 1.013 | 0.994 | 1.033 |

Table 7.5.1d The differences between groups after intervention: amoxycillin prescriptions.

| Between group comparison after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 0.3% | 0.15 | 1.015 | 0.995 | 1.037 |
| C *versus* D | 1.7% | <0.0001 | 1.115 | 1.093 | 1.138 |
| S *versus* D | 1.4% | <0.0001 | 1.098 | 1.077 | 1.12 |

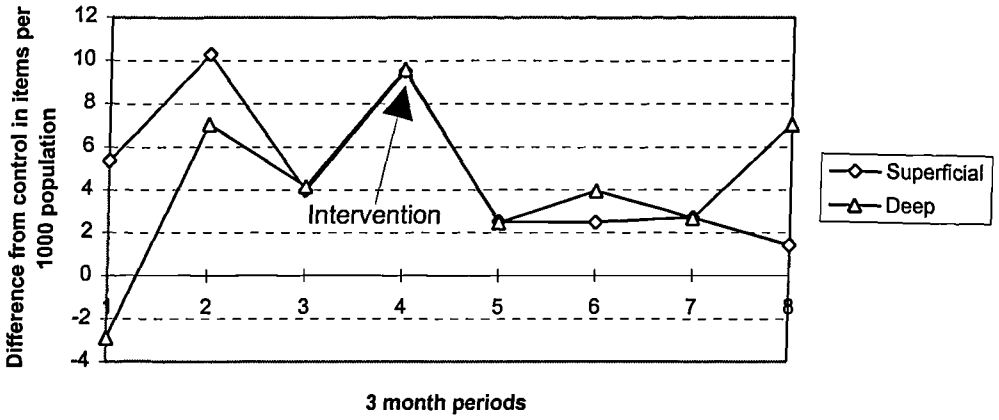



Figure 7.5.1a The difference in items of amoxycillin prescribed per 1000 population per quarter between the intervention groups and control (i.e. zero represents control).

### 7.5.2 Co-amoxiclav

The guidelines recommended co-amoxiclav for sinusitis, pneumonia, bites and unobstructed cholecystitis. All groups became significantly more likely to prescribe co-amoxiclav after intervention. There was no significant difference between groups before intervention. Afterwards, however, the superficial group was significantly more likely to prescribe co-amoxiclav than either control or deep.

The superficial groups' large response to recommended co-amoxiclav in the first quarter (figure) explains the significantly favourable difference from control and deep in the twelve months post-intervention. The effect is dramatic, but does not persist.

Table 7.5.2a Co-amoxiclav prescriptions as proportions of total prescriptions in the year before and year after intervention.

| | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| Antibiotic. | before | after | before | after | before | after |
| Total items | 125897 | 114549 | 134920 | 125479 | 145369 | 134884 |
| Co-amoxiclav | 7.5% | 9.7% | 7.4% | 10.2% | 7.4% | 9.8% |

Table 7.5.2b The difference between before and after intervention in each group: co-amoxiclav prescriptions.

| Within group comparison before *versus* After (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | 2.2% | <0.0001 | 1.318 | 1.281 | 1.357 |
| Superficial | 2.8% | <0.0001 | 1.418 | 1.38 | 1.458 |
| Deep | 2.4% | <0.0001 | 1.362 | 1.326 | 1.399 |

Table 7.5.2c The differences between groups before intervention: co-amoxiclav prescriptions.

| Between group comparison before (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | -0.1% | 0.48 | 1.011 | 0.981 | 1.041 |
| C *versus* D | -0.1% | 0.16 | 1.021 | 0.992 | 1.051 |
| S *versus* D | -0.1% | 0.49 | 1.01 | 0.982 | 1.039 |

Table 7.5.2d The differences between groups after intervention: co-amoxiclav prescriptions.

| Between group comparison after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 0.6% | <0.0001 | 1.065 | 1.036 | 1.094 |
| C *versus* D | 0.1% | 0.38 | 1.012 | 0.986 | 1.04 |
| S *versus* D | -0.5% | <0.0001 | 1.052 | 1.025 | 1.079 |

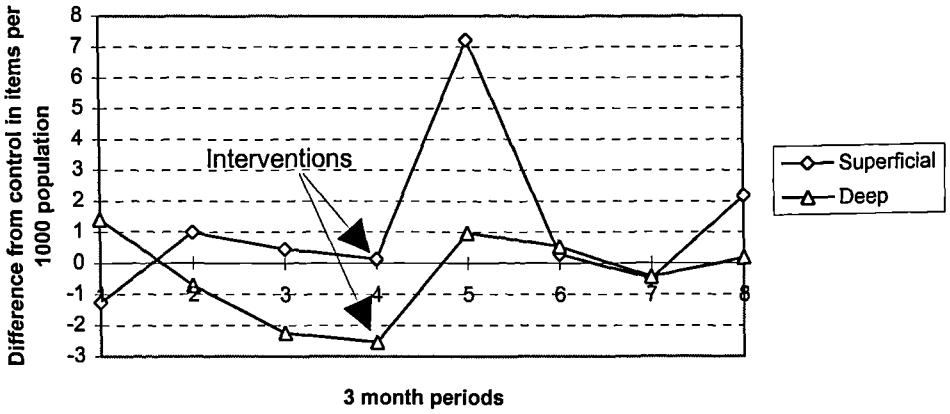



Figure 7.5.2a The difference in items of co-amoxiclav prescribed per 1000 population per quarter between the intervention groups and control (i.e. control = zero). Note: the large 'novelty' effect in the superficial group in the first 3 months post-intervention (period 5) which is not sustained.

### 7.5.3 Trimethoprim.

Trimethoprim is the recommended antibiotic for cystitis in the guidelines. Before intervention superficial and deep groups were significantly more likely to prescribe trimethoprim than control, but there was no significant difference between them. All three groups became significantly more likely to prescribe trimethoprim after intervention. After intervention the deep group had become statistically highly significantly more likely than the superficial to prescribe trimethoprim.

Table 7.5.3a Trimethoprim prescriptions as proportions of total prescriptions in the year before and year after intervention.

| | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| Antibiotic. | before | after | before | after | before | after |
| Trimethoprim | 5911 | 5647 | 7283 | 7240 | 7966 | 8538 |
| Total items | 125897 | 114549 | 134920 | 125479 | 145369 | 134884 |
| % of total. | 4.7 | 4.9 | 5.4 | 5.8 | 5.5 | 6.3 |

Table 7.5.3b The difference between before and after intervention in each group: Trimethoprim prescriptions.

| Within group comparison before *versus* After (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | 0.3% | 0.002 | 1.061 | 1.022 | 1.102 |
| Superficial | 0.4% | <0.0001 | 1.073 | 1.038 | 1.12 |
| Deep | 0.9% | <0.0001 | 1.166 | 1.203 | 1.203 |

Table 7.5.3c The differences between groups before intervention: Trimethoprim prescriptions.

| Between group comparison before (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 0.7% | <0.0001 | 1.168 | 1.127 | 1.21 |
| C *versus* D | 0.8% | <0.0001 | 1.187 | 1.146 | 1.229 |
| S *versus* D | 0.1% | 0.34 | 1.016 | 0.983 | 1.05 |

Table 7.5.33d The differences between groups after intervention: Trimethoprim prescriptions.

| Between group comparison after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 0.8% | <0.0001 | 1.181 | 1.139 | 1.224 |
| C *versus* D | 1.4% | <0.0001 | 1.303 | 1.259 | 1.349 |
| S *versus* D | 0.6% | <0.0001 | 1.104 | 1.068 | 1.14 |

### 7.5.4 Guideline antibiotics.

The groups differ significantly from each other and from themselves in their likelihood to prescribe guideline antibiotics both before and after intervention. However, superficial was the most likely group to prescribe guideline antibiotics before intervention, the deep after.

Table 7.5.4a Guideline antibiotic prescriptions as proportions of total prescriptions in the year before and year after intervention.

| | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| Antibiotic. | before | after | before | after | before | after |
| Total items | 125897 | 114549 | 134920 | 125479 | 145369 | 134884 |
| Guideline antibiotics. | 45.5% | 50.6% | 49.2% | 51.4% | 49.1% | 52.3% |

Table 7.5.4b The difference between before and after intervention in each group: guideline antibiotic prescriptions.

| Within group comparison before *versus* After (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | 5.1% | <0.0001 | 1.227 | 1.207 | 1.247 |
| Superficial | 3.3% | <0.0001 | 1.125 | 1.108 | 1.143 |
| Deep | 3.3% | <0.0001 | 1.138 | 1.122 | 1.155 |

Table 7.5.4c The differences between groups before intervention: guideline antibiotic prescriptions.

| Between group comparison before (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 3.7% | <0.0001 | 1.129 | 1.111 | 1.146 |
| C *versus* D | 3.6% | <0.0001 | 1.155 | 1.138 | 1.173 |
| S *versus* D | -0.1% | 0.002 | 1.023 | 1.008 | 1.039 |

Table 7.5.4d The differences between groups after intervention: guideline antibiotic prescriptions.

| Between group comparison after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 0.9% | <0.0001 | 1.035 | 1.019 | 1.052 |
| C *versus* D | 1.7% | <0.0001 | 1.072 | 1.055 | 1.089 |
| S *versus* D | 0.9% | <0.0001 | 1.036 | 1.02 | 1.052 |

### 7.5.5 Newer antibiotics.

The guidelines did not promote any of the newer (commercially promoted) antibiotics. For purpose of analysis the 'newer drugs' selected were clarithromycin, cefixime, ciprofloxacin, ofloxacin and norfloxacin (as these were relatively common and heavily promoted). The three groups differ in their likelihood to prescribe newer antibiotics both before and after intervention. There have been significant changes in all groups. However, the control and superficial have significantly increased likelihood to prescribe newer antibiotics, the deep significantly less than previously.

Table 7.5.5a Newer antibiotic prescriptions as proportions of total prescriptions in the year before and year after intervention.

| | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| Antibiotic. | before | after | before | after | before | after |
| Total items | 125897 | 114549 | 134920 | 125479 | 145369 | 134884 |
| Newer antibiotics | 3.6% | 4.3% | 3.5% | 4% | 4.1% | 3.9% |

Table 7.5.5b The difference between before and after intervention in each group: newer antibiotic prescriptions.

| Within group comparison before *versus* After (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | 0.7% | <0.0001 | 1.214 | 1.165 | 1.266 |
| Superficial | 0.5% | <0.0001 | 1.148 | 1.102 | 1.196 |
| Deep | -0.2% | 0.04 | 1.04 | 1.001 | 1.08 |

Table 7.5.5c The differences between groups before intervention: newer antibiotic prescriptions.

| Between group comparison before (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | -0.1% | 0.13 | 1.033 | 0.991 | 1.077 |
| C *versus* D | 0.5% | <0.0001 | 1.134 | 1.09 | 1.18 |
| S *versus* D | 0.6% | <0.0001 | 1/171 | 1.126 | 1.218 |

Table 7.5.5d The differences between groups after intervention: newer antibiotic prescriptions.

| Between group comparison after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | -0.4% | <0.0001 | 1.093 | 1.049 | 1.138 |
| C *versus* D | -0.4% | <0.0001 | 1.113 | 1.07 | 1.159 |
| S *versus* D | -0.1% | 0.04 | 1.019 | 0.979 | 1.06 |

### 7.5.6 Co-trimoxazole.

Trimethoprim may be prescribed on its own or in combination with a sulphonamide (co-trimoxazole); the indications are identical. The lectures and visits recommended that trimethoprim be substituted for co-trimoxazole. All three groups were significantly different in their likelihood to prescribe co-trimoxazole rather than trimethoprim both before and after intervention. Each group had a reduced tendency to prescribe co-trimoxazole after intervention. However, the deep group showed the biggest change.

Table 7.5.6a Trimethoprim and co-trimoxazole prescriptions in the four (3 month) quarters before and after intervention.

| | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| Antibiotic. | before | after | before | after | before | after |
| Trimethoprim | 5911 | 5647 | 7283 | 7240 | 7966 | 8538 |
| Co-trimoxazole | 3352 | 2043 | 4379 | 3006 | 5343 | 3236 |
| Percentage co-trim. | 36.2% | 26.6% | 37.6% | 29.3% | 40.2% | 27.5% |

Table 7.5.6b The difference between before and after intervention in each group: proportion of trimethoprim: proportion of trimethoprim as co-trimoxazole.

| Within group comparison before *versus* after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | -9.6% | <0.0001 | 1.567 | 1.467 | 1.675 |
| Superficial | -8.2% | <0.0001 | 1.448 | 1.368 | 1.533 |
| Deep | -12.7% | <0.0001 | 1.77 | 1.677 | 1.867 |

Table 7.5.6c The differences between groups before intervention: proportion of trimethoprim as co-trimoxazole.

| Between group comparison before (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 1.4% | 0.044 | 1.06 | 1.002 | 1.122 |
| C *versus* D | 4.0% | <0.0001 | 1.183 | 1.119 | 1.25 |
| S *versus* D | 2.6% | <0.0001 | 1.116 | 1.06 | 1.174 |

Table 7.5.6d The differences between groups after intervention: proportion of trimethoprim as co-trimoxazole.

| Between group comparison after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 2.8% | <0.0001 | 1.148 | 1.074 | 1.227 |
| C *versus* D | 0.9% | 0.16 | 1.048 | 0.981 | 1.118 |
| S *versus* D | -1.9% | 0.002 | 1.095 | 1.033 | 1.162 |

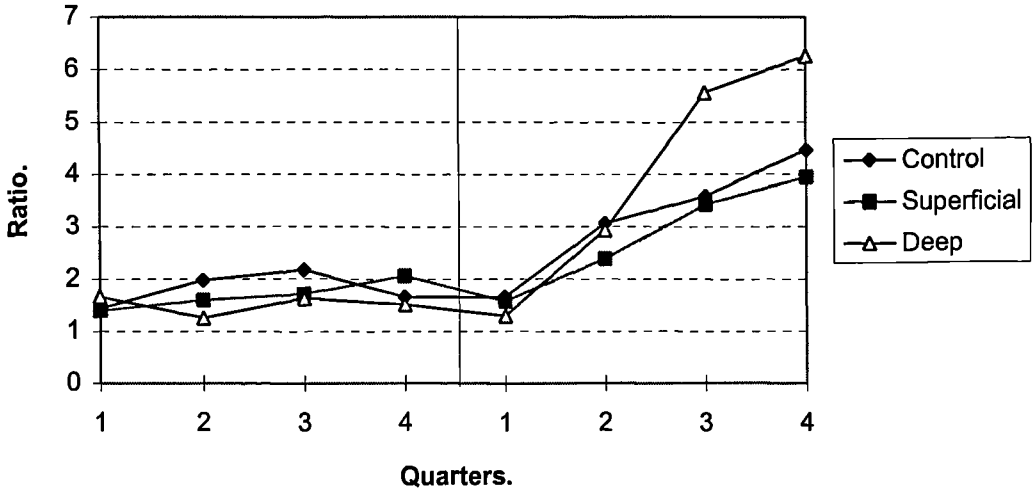



Figure 7.5.6a - The ratio of trimethoprim to co-trimoxazole prescriptions in the four (3 month) quarters before and after intervention.

### 7.5.7 Tetracyclines.

Tetracyclines do not figure at all in the guidelines and were discouraged in both lectures and visits. Only the deep group had a statistically significantly reduced likelihood of prescribing tetracyclines after intervention.

Table 7.5.7a Tetracycline prescriptions as proportions of total prescriptions in the year before and year after intervention.

| | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| Antibiotic. | before | after | before | after | before | after |
| Total items | 125897 | 114549 | 134920 | 125479 | 145369 | 134884 |
| Tetracyclines | 6% | 5.8% | 6% | 6% | 5.6% | 5.1% |

Table 7.5.7b The difference between before and after intervention in each group: tetracycline prescriptions.

| Within group comparison before *versus* After (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | -0.16% | 0.1 | 1.029 | 0.994 | 1.064 |
| Superficial | 0.03% | 0.7 | 1.007 | 0.974 | 1.04 |
| Deep | -0.42% | <0.0001 | 1.087 | 1.051 | 1.123 |

Table 7.5.7c The differences between groups before intervention: tetracycline prescriptions.

| Between group comparison before (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 0% | 0.99 | 1 | 0.968 | 1.033 |
| C *versus* D | -0.44% | <0.0001 | 1.084 | 1.049 | 1.12 |
| S *versus* D | -0.44% | <0.0001 | 1.084 | 1.049 | 1.119 |

Table 7.5.7d The differences between groups after intervention: tetracycline prescriptions.

| Between group comparison after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 0.19% | 0.047 | 1.035 | 1 | 1.071 |
| C *versus* D | -0.7% | <0.0001 | 1.145 | 1.106 | 1.186 |
| S *versus* D | 0.89% | <0.0001 | 1.185 | 1.146 | 1.226 |

### 7.5.8 Generic Prescribing

The guidelines employed generic antibiotic names exclusively. All groups demonstrate significant improvements in generic prescribing although greatest in the deep. The national generic prescribing percentage in 1993 was 64.9%, and in 1994 was 72.0%; an increase of 7.1%. The control and deep groups both exceeded the national change perhaps reflecting the success of local initiatives to increase generic prescribing generally.

Table 7.5.8a Generic prescriptions as proportions of total prescriptions in the year before and year after intervention.

| | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| | before | after | before | after | before | after |
| Total items | 125897 | 114549 | 134920 | 125479 | 145369 | 134884 |
| Generic items | 48.9% | 57% | 55.2% | 61.5% | 51% | 59.9% |

Table 7.5.8b The difference between before and after intervention in each group: generic prescriptions.

| Within group comparison before *versus* After (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | 8.1% | <0.0001 | 1.1653 | 1.1495 | 1.1813 |
| Superficial | 6.3% | <0.0001 | 1.1143 | 1.1002 | 1.1285 |
| Deep | 9.0% | <0.0001 | 1.1760 | 1.1615 | 1.907 |

Table 7.5.8c The differences between groups before intervention: generic prescriptions.

| Between group comparison before (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 6.3% | <0.0001 | 1.1293 | 1.1145 | 1.1442 |
| C *versus* D | 2.1% | <0.0001 | 1.0422 | 1.0287 | 1.056 |
| S *versus* D | -4.3% | <0.0001 | 1.0835 | 1.0699 | 1.0972 |

Table 7.5.8d The differences between groups after intervention: generic prescriptions.

| Between group comparison after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 4.6% | <0.0001 | 1.0798 | 1.0657 | 1.0941 |
| C *versus* D | 3.0% | <0.0001 | 1.0518 | 1.0383 | 1.0656 |
| S *versus* D | -1.6% | <0.0001 | 1.0266 | 1.0138 | 1.0395 |

## 7.5.9 Oral antibiotic prescribing patterns in the final 3 months.

In the final study quarter (9 to 12 months post intervention) the numbers of items prescribed of oral antibiotics were as follows, ranked according to popularity in the control group:

Table 7.5.9a Actual oral antibiotic prescriptions in each group in the final 3 months of the study (9 to 12 months post-intervention) ranked according to control volume.

| Antibiotic | Control | Superficial | Deep |
|---|---|---|---|
| Amoxycillin | 4694 | 5115 | 6507 |
| Co-amoxiclav | 2764 | 3182 | 3263 |
| Penicillin V | 2646 | 1805 | 1774 |
| Erythromycin | 1736 | 2356 | 2024 |
| Trimethoprim | 1628 | 2011 | 2345 |
| Cephalexin | 982 | 1577 | 1790 |
| Oxytetracycline | 705 | 844 | 753 |
| Flucloxacillin | 705 | 758 | 1089 |
| Doxycycline | 621 | 739 | 665 |
| Cefaclor | 601 | 729 | 778 |
| Minocycline | 494 | 707 | 644 |
| Ampicillin | 414 | 166 | 43 |
| Ciprofloxacin | 388 | 603 | 694 |
| Cephradine | 381 | 111 | 243 |
| Co-trimoxazole | 362 | 502 | 355 |
| Metronidazole | 351 | 438 | 442 |
| Clarithromycin | 287 | 197 | 449 |
| Cefixime | 256 | 299 | 80 |
| Co-fluampicil | 198 | 186 | 136 |
| Ofloxacin | 193 | 175 | 174 |
| Nitrofurantoin | 190 | 147 | 208 |
| Tetracycline | 103 | 194 | 155 |
| Norfloxacin | 36 | 5 | 12 |
| Cefuroxime | 15 | 38 | 23 |
| Total | 20750 | 22884 | 24646 |

## Penicillins

The penicillin prescriptions in the last 3 months of the study were:

Table 7.5.9b Prescriptions for penicillins in the final 3 months of the study (9 to 12 months post-intervention).

| Penicillins | Control | Superficial | Deep |
|---|---|---|---|
| Amoxycillin | 4694 | 5115 | 6507 |
| Co-amoxiclav | 2764 | 3182 | 3263 |
| Penicillin V | 2646 | 1805 | 1774 |
| Flucloxacillin | 705 | 758 | 1089 |
| Ampicillin | 414 | 166 | 43 |
| Co-fluampicil | 198 | 186 | 136 |
| Total | 11421 | 11212 | 12812 |

The guidelines recommended amoxycillin, co-amoxiclav, penicillin V, flucloxacillin. After 9 months superficial group GPs were significantly more likely to prescribe guideline penicillins than control. However, the deep group GPs were significantly more likely to prescribe them than both control and superficial. The proportion of the above penicillins in the guidelines were:

Table 7.5.9c Guideline penicillin prescriptions as a proportion of total penicillin prescriptions.

| Group | Control | Superficial | Deep |
|---|---|---|---|
| Penicillins in guidelines | 94.6% | 96.9% | 98.6% |

Table 7.5.9d The differences between groups in likelihood to use guideline penicillins 9 to 12 months after intervention.

| Between group comparison 9 to 12 months after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 2.2% | <0.0001 | 1.7458 | 1.5247 | 2.0015 |
| C *versus* D | 4.0% | <0.0001 | 3.988 | 3.365 | 4.7468 |
| S *versus* D | 1.7% | <0.0001 | 2.2844 | 1.8999 | 2.755 |

## Cephalosporins

The prescriptions for the oral cephalosporins in the final 3 months of the study year were as follows:

Table 7.5.9e Prescriptions for cephalosporins in the final 3 months of the study (9 to 12 months post-intervention).

| Cephalosporin | Control | Superficial | Deep |
|---|---|---|---|
| Cephalexin | 982 | 1577 | 1790 |
| Cefaclor | 601 | 729 | 778 |
| Cephradine | 381 | 111 | 243 |
| Cefixime | 256 | 299 | 80 |
| Cefuroxime | 15 | 38 | 23 |
| Total | 2235 | 2754 | 2914 |

The guidelines recommended only cephalexin. The superficial group was significantly more likely to prescribe cephalexin than the control, however the deep group was more likely than either superficial or control.

Table 7.5.9f The differences between groups in likelihood to use the guideline cephalosporin 9 to 12 months after intervention.

| Between group comparison 9 to 12 months after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 13.3% | <0.0001 | 1.709 | 1.525 | 1.916 |
| C *versus* D | 17.5% | <0.0001 | 2.031 | 1.814 | 2.277 |
| S *versus* D | 4.2% | 0.002 | 1.189 | 1.067 | 1.324 |

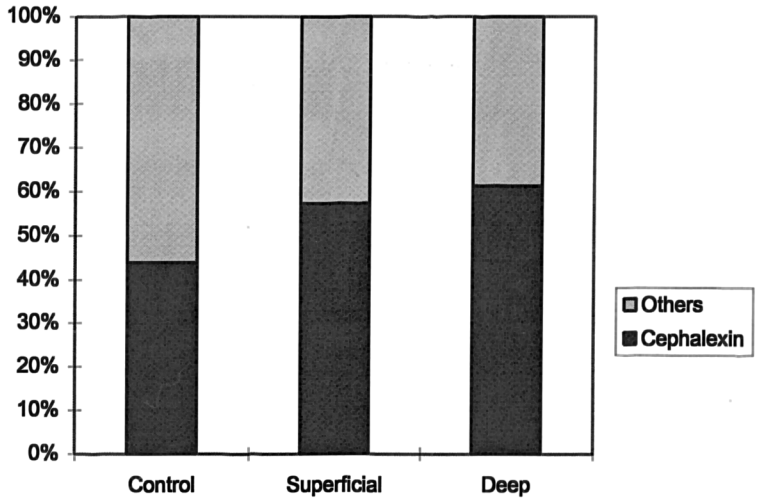



Figure 7.5.9a - The proportion of cephalexin versus other oral cephalosporins prescribed by the three groups in August to October 1994 (the final quarter of the study year).

### 7.5.10 Prescribing cost and volume.

## Change in prescribing cost

If antibiotic prescribing is more efficient then the cost per prescription item should fall.

Table 7.5.10a Cost per antibiotic prescription in the year before and year after intervention (not discounted).

| | Control | | Superficial | | Deep | |
|---|---|---|---|---|---|---|
| | before | after | before | after | before | after |
| Cost per item | £4.13 | £4.20 | £4.09 | £4.21 | £4.09 | £4.13 |
| Increase | - | 6.9p | - | 11.6p | - | 3.7p |

The cost increase in the deep group was only 53% of that expected from control, whereas the superficial increase was 69% more. The post-intervention difference in cost per item in the deep group relative to control represents a financial saving of £9,500 (7p per prescription) or 1.6% on antibiotic prescribing costs. Greater cost in the superficial group may have been caused by an increasing tendency to prescribe new drugs (see 'new drugs' above). The elaboration likelihood model is consistent with this: it predicts susceptible attitudes after superficial processing. In England as a whole the cost per antibiotic prescription in 1993 was £4.20, and in 1994, £4.47; an increase of 27p (7%).

## Change in prescribing rate.

Rates per head of population depend upon FHSA registration data which is of variable accuracy, so caution is required in interpretation. Prescribing rates of antibiotics drugs before and after the interventions were as follows:

Table 7.5.10b Antibiotic prescriptions issued per 1000 population in the year before and the year after intervention.

| | **Control** | | **Superficial** | | **Deep** | |
|---|---|---|---|---|---|---|
| | before | after | before | after | before | after |
| Population | 121143 | 121519 | 128142 | 127883 | 144733 | 142679 |
| Items | 125897 | 114549 | 134920 | 125479 | 145369 | 134884 |
| Rate per 1000 population. | 1044.21 | 942.64 | 1052.89 | 981.2 | 1004.39 | 945.37 |

All England antibiotic prescribing rates fell by 3% over the same period.

Prescribing costs for antibiotics were:

Table 7.5.10c Actual antibiotic cost per person in the year before and year after intervention (Drug Tariff prices).

| | Control | | Superficial | | Deep | |
|---|---|---|---|---|---|---|
| | before | after | before | after | before | after |
| Population | 121143 | 121519 | 128142 | 127883 | 144733 | 142679 |
| Cost £ | £519,841 | £480,880 | £551,857 | £527,855 | £595,082 | £557,083 |
| Cost per person | £4.29 | £3.96 | £4.31 | £4.13 | £4.11 | £3.9 |

The deep group is in a slightly more favourable position with regard to cost per patient after the intervention than the other groups. Rate of prescribing is similar in both control and deep. However, the superficial group is in a less favourable position on both volume and cost. The rate of prescribing has reduced in all groups perhaps reflecting a lower incidence of infection in 1994. All England antibiotic prescribing expenditure increased by 3% over the same period.

## Summary of prescribing behaviour.

Time and again more favourable prescribing behaviour is demonstrated in the intervention groups. Generally, the difference from the control is more favourable in deep than superficial. There is evidence of persistence in the deep group. Conversely, there is evidence of susceptibility to commercial counter-persuasion in the superficial group (i.e. proportion of new drugs). As anticipated the rate of prescribing has not been affected, but there have been qualitative changes in prescribing in the deep group including a reduction in cost per head relative to the other groups.

Table 7.5z Summary statistics of changes in prescribing indicators after interventions relative to control group changes. Where reductions have been advocated the polarity has been reversed before analysis.

| Group | Mean | 95% CL | Median | Minimum to maximum |
|---|---|---|---|---|
| Superficial | -0.9% | -1.69 to -0.05 | -1.1% | -2.4 to 0.6 |
| Deep | 0.6% | -0.5 to 1.6 | 0.6% | -1.8 to 3.1 |

## 7.6 Laboratory Investigations

Laboratory investigations increased by about 25% in all three groups over the study period. Most types of investigation are too rare and sporadic for meaningful analysis, therefore presented below are analyses of the three commonest types of investigation from general practice.

### 7.6.1 Investigations for urinary tract infection.

The investigations are chiefly mid-stream specimens of urine (MSSU). The guidelines recommended empirical treatment of cystitis, reserving investigation for complicated or persistent cases. The greatest change post-intervention is in the deep processing group and represents a reduction in investigation. This is consistent with the guideline recommendations for empirical treatment of uncomplicated cystitis.

Table 7.6.1a Investigations for urinary tract infections (U.T.I.) as proportions of total microbiological investigations in the year before and year after intervention.

| | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| | before | after | before | after | before | after |
| Total tests | 10338 | 12654 | 11037 | 14047 | 12872 | 15952 |
| U.T.I. tests | 58.6% | 54.4% | 57.9% | 52.9% | 61.4% | 54% |

Table 7.6.1b The difference between before and after intervention in each group: investigations for urinary tract infections.

| Within group comparison before *versus* after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | -4.2% | <0.0001 | 1.1851 | 1.1241 | 1.2495 |
| Superficial | -5.0% | <0.0001 | 1.2254 | 1.1649 | 1.289 |
| Deep | -7.4% | <0.0001 | 1.3552 | 1.2924 | 1.4212 |

Table 7.6.1c The differences between groups before intervention: investigations for urinary tract infections.

| Between group comparison before (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | -0.62% | 0.37 | 1.0258 | 0.9711 | 1.0835 |
| C *versus* D | 2.82% | <0.0001 | 1.1248 | 1.0665 | 1.1862 |
| S *versus* D | 3.44% | <0.0001 | 1.1538 | 1.0951 | 1.2156 |

Table 7.6.1d The differences between groups after intervention: investigations for urinary tract infections.

| Between group comparison after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | -1.46% | 0.0173 | 1.0606 | 1.0104 | 1.1133 |
| C *versus* D | -0.41% | 0.5 | 1.0166 | 0.9699 | 1.0657 |
| S *versus* D | 1.05% | 0.07 | 1.0432 | 0.9966 | 1.092 |

### 7.6.2 Investigations for gynaecological infections

The chief investigations are vaginal swabs. The guidelines recommended referral of cases of pelvic inflammatory disease to genito-urinary specialist or gynaecologist for full investigation before treatment is exhibited. The deep processing group did not change significantly, however, the recommendation in the guidelines to refer all cases of pelvic infection was said to be impractical by several of the visited general practitioners.

Table 7.6.2a Investigations for gynaecological infections as proportions of total microbiological investigations in the year before and year after intervention.

| | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| | before | after | before | after | before | after |
| Total tests | 10338 | 12654 | 11037 | 14047 | 12872 | 15952 |
| Gynae. tests | 16.3% | 14.2% | 16% | 14.6% | 11.2% | 10.6% |

Table 7.6.2b The difference between before and after intervention in each group: investigations for gynaecological infections.

| Within group comparison before *versus* after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | -2% | <0.0001 | 1.171 | 1.0887 | 1.2595 |
| Superficial | -1.4% | 0.002 | 1.1164 | 1.041 | 1.1971 |
| Deep | -0.6% | 0.1 | 1.0656 | 0.9883 | 1.1489 |

Table 7.6.2c The differences between groups before intervention: investigations for gynaecological infections.

| Between group comparison before (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | -0.3% | 0.57 | 1.0222 | 0.9496 | 1.1003 |
| C *versus* D | -5.1% | <0.0001 | 1.54497 | 1.4313 | 1.6679 |
| S *versus* D | -4.8% | <0.0001 | 1.5114 | 1.4016 | 1.6302 |

Table 7.6.2d The differences between groups after intervention: investigations for gynaecological infections.

| Between group comparison after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | 0.3% | 0.47 | 1.0262 | 0.9577 | 1.0996 |
| C *versus* D | -3.7% | <0.0001 | 1.406 | 1.3088 | 1.5104 |
| S *versus* D | -4% | <0.0001 | 1.4427 | 1.3459 | 1.5467 |

### 7.6.3 Investigations for intestinal infections

The guidelines recommended investigation of all cases of gastro-enteritis persisting for three days. Only the deep group showed significant evidence of increased rate of investigation.

Table 7.6.3a Investigations for intestinal infections as proportions of total microbiological investigations in the year before and year after intervention.

| | Control: | | Superficial: | | Deep: | |
|---|---|---|---|---|---|---|
| | before | after | before | after | before | after |
| Total tests | 10338 | 12654 | 11037 | 14047 | 12872 | 15952 |
| Faecal tests | 7.9% | 7.9% | 7% | 7.4% | 8.3% | 9.1% |

Table 7.6.3b The difference between before and after intervention in each group: investigations for intestinal infections.

| Within group comparison before *versus* after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| Control | 0% | 1 | 1 | 0.9073 | 1.1027 |
| Superficial | 0.4% | 0.19 | 1.0685 | 0.9691 | 1.1784 |
| Deep | 0.8% | 0.02 | 1.1041 | 1.0157 | 1.2005 |

Table 7.6.3c The differences between groups before intervention: investigations for intestinal infections.

| Between group comparison before (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | -0.9% | 0.01 | 1.141 | 1.0287 | 1.2657 |
| C *versus* D | 0.41% | 0.27 | 1.0557 | 0.959 | 1.1624 |
| S *versus* D | 1.31% | 0.0002 | 1.2045 | 1.0928 | 1.3281 |

Table 7.6.3d The differences between groups after intervention: investigations for intestinal infections.

| Between group comparison after (Chi squared test): | | | | | |
|---|---|---|---|---|---|
| | Difference | P | Odds ratio | Lower CL | Upper CL |
| C *versus* S | -0.46% | 0.16 | 1.068 | 0.9748 | 1.1701 |
| C *versus* D | 1.19% | 0.0004 | 1.1655 | 1.0705 | 1.2692 |
| S *versus* D | 1.65% | <0.0001 | 1.2447 | 1.1447 | 1.3538 |

## Summary of laboratory investigation changes

The empirical antibiotic recommendations for cystitis in the guidelines is based upon community-acquired infection data; provided that the recommended antibiotic is chosen initial investigation is unnecessary. Consequently deep processing should result in less investigation. The deep group sent significantly less MSSU investigations following intervention than the other groups.

The guidelines recommendation to refer all cases of pelvic infection for thorough investigation resulted in significantly less reduction in the proportion of vaginal swab investigations than occurred in the other groups. However, several in the deep processing group felt that this recommendation was impractical.

The guidelines recommended investigation of all cases of gastro-enteritis unresolved at 3 days. The deep group increased its proportion of faecal investigations significantly more than the other groups post-intervention.

The same effects were exhibited by the superficial group relative to control, but of significantly less magnitudes.

Table 7.6z Summary statistics of changes in laboratory investigation indicators after interventions relative to control group changes (3 indicators). Where reductions have been advocated the polarity has been reversed before analysis.

| Group | Mean | 95% CL | Median | Minimum to maximum |
|---|---|---|---|---|
| Superficial | 0.6% | 0.1 to 1.1 | 0.6% | 0.4 to 0.8 |
| Deep | 1.8% | -1.3 to 4.9 | 1.4% | 0.8 to 3.2 |

## 7.7 Clinical Study

At the end of the study year a survey of cystitis management was performed prospectively for one month by Wirral GPs from each group. The guidelines recommended trimethoprim or cephalexin for simple cystitis:

Table 7.7a Percentage prescriptions for guideline antibiotics in the cystitis study.

| | Control | Superficial | Deep |
|---|---|---|---|
| n = | 90 | 41 | 49 |
| Guideline antibiotics | 85.6% | 51.2% | 87.8% |

These results echo the attitude questionnaire and prescribing results above. Trimethoprim use increases with depth of processing. The use of the new antibiotic norfloxacin is absent from the deep group, but is used by the control and, most heavily, the superficial:

Table 7.7b Norfloxacin prescriptions as a percentage of total antibiotic prescriptions in the cystitis study.

| | Control | Superficial | Deep |
|---|---|---|---|
| n = | 90 | 41 | 49 |
| Norfloxacin | 2.2% | 4.9% | 0% |

There was no general difference in length of course; the mode is 5 days in all groups.

Table 7.7c The distribution of antibiotic length of course in the cystitis study (modes in bold type)

| Length of course | Control | Superficial | Deep |
|---|---|---|---|
| 0 days | 2 | 2 | 0 |
| 3 days | 24 | 15 | 8 |
| **5 days** | **38** | **17** | **30** |
| 7 days | 26 | 4 | 9 |
| Other | 0 | 3 | 2 |
| *n* = | *90* | *41* | *49* |

The rate of mid-stream specimen of urine investigation was lower in the superficial group:

Table 7.7d The percentage of cases investigated (mid-stream specimen of urine) by referral of specimen to laboratory in the cystitis study.

| | Control | Superficial | Deep |
|---|---|---|---|
| n = | 90 | 41 | 49 |
| MSSU | 56% | 46% | 57% |

The number of cases in the audit provides insufficient statistical power for hypothesis testing of the small differences in prescribing and laboratory investigation. However, the purpose of the audit was to explore patient outcome; something the other elements of the study were unable to do. The principal outcome measure of successful treatment in a curable disease is absence of relapse; there were no relapses in the deep group. The differences between the deep and control, and deep and superficial are both statistically significant. It is unlikely that the differences have occurred by chance.

7.7e Relapse / return rate in two week period after initial consultation in the cystitis study.

| | Control | Superficial | Deep |
|---|---|---|---|
| n = | 90 | 41 | 49 |
| Relapse rate | 8 | 6 | 0 |

Table 7.7f - Analysis (of unpaired proportions) of relapse in the three groups.

| | Difference in proportions | P | Upper CL | Lower CL |
|---|---|---|---|---|
| Control *versus* Superficial | 5.8% | 0.3237 | -0.204 | 0.052 |
| Control *versus* Deep | 8.9% | 0.03 | -0.014 | 0.166 |
| Superficial *versus* Deep | 14.6% | 0.0056 | 0.068 | 0.284 |

Application of Hanley's formula*[205] to the above suggests that the rate of re-attendance in the patients of doctors who have deeply processed cystitis issues is at the very most 6% rather than the expected 9%.

*Hanley's rule: If none of n patients showed the event about which we are concerned, we can be 95% confident that the chance of this event is at most 3/n - the upper 95% CL of a zero/n rate is 3/n.

## 7.8 Health Economics

This section examines the costs and relative cost-effectiveness of the superficial and deep approaches. As the population figures (drawn from Family Health Service Authority registers) had proven unreliable due to variable capture, administrative problems, and list inflation ('ghost patients')[206] effectiveness is expressed as either cost per unit change per GP, or cost per unit change per prescription. The number of GPs remained constant in each group throughout, and the number of prescriptions dispensed is collected by the Prescribing Pricing Authority with 99.9% accuracy.[207] Ascribing a value to many of the indicators is contentious and the figures are offered to provoke thought rather than to establish proof.

### Cost per general practitioner

The denominator most appropriate is the number of GPs as they were the 'subjects' of the experiment.

Indirect costs are assumed to be zero as GPs are obliged to attend approved continuing education events. The cost of preparation and production of the guidelines was split equally between both intervention groups. The direct costs for the

superficial group included lecturer's fees, administration and organisation of the event, and accreditation fee.

The direct costs of the deep group were calculated as one third of a consultant session for two consultants per visit, travelling costs at an average 12 miles per visit at 35p per mile.

Table 7.8a The cost of the interventions.

| | Control | Superficial | Deep |
|---|---|---|---|
| Cost | £0 | £1,909 | £5060 |
| No. Of GPs | 55 | 63 | 64 |
| Cost per GP | £0 | £30.30 | £79.06 |

## Prevention of relapses in cystitis

The cystitis study showed a significant reduction in relapse (re-attendance with the cystitis within 2 weeks of first consultation) in the deep group. The expected cases are calculated from the 4th National Morbidity Survey.[208]

Table 7.8b The cost per prevented relapse in each group.

| | Control | Superficial | Deep |
|---|---|---|---|
| Cases expected | 1337 | 1406 | 1569 |
| Relapses expected | 118 | 125 | 140 |
| Actual relapses | 118 | 205 | 0 |
| Prevented relapses | 0 | -80 | 140 |
| Cost per prevented relapse | n.a. | n.a. | £36.14 |

## 7.9 Effect on individual practices

Most individual practice level effects are obscured by confounding statistical 'noise', however there are examples of small stable practices where effect is observable; they are displayed below to illustrate the kind type of phenomenon that depth of processing theory predicts. The effect on cost of antibiotics and proportion prescribed generically is more pronounced in the examples than seen in the aggregated data:

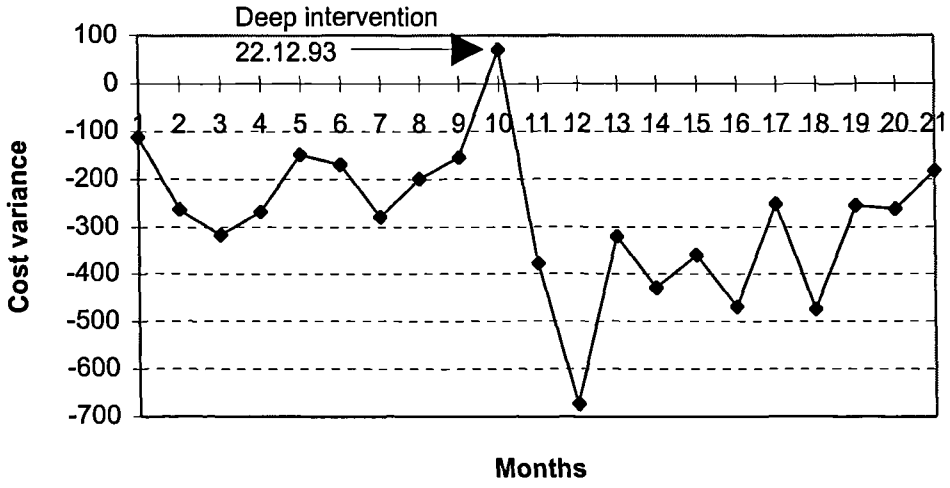



Figure 7.9a The effect of deep processing on a single-handed practice: Difference ('variance') in cost of antibiotic prescriptions from that predicted from the district cost per head

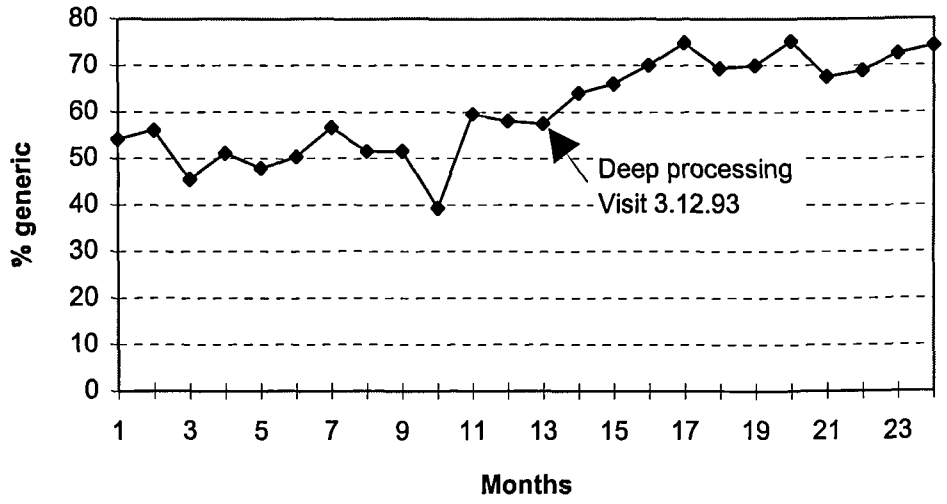



Figure 7.9b The effect of deep processing on a four-handed practice: the percentage of antibiotic prescriptions written generically.

Both practices demonstrate persistent changes after the deep processing intervention. A practice in the superficial group by contrast demonstrates the classical short-lived effect of superficial processing:

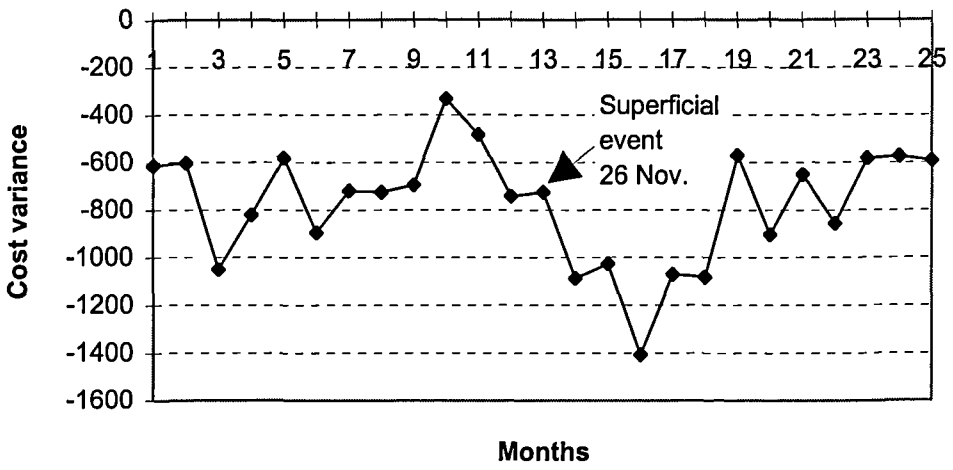



Figure 7.9c The effect of superficial processing on a three-handed practice: difference in cost ('variance') of antibiotic prescriptions from that expected by district average cost per head. Note the downward effect is less sustained than in figure 7.4.

No practice in the superficial group displayed pronounced change in generic antibiotic prescribing.

# Chapter 8

# Consequences of depth of processing in the study

*A controversy prevailed among the beasts of the field, as to which of the animals deserved the most credit for producing the most whelps at birth. They rushed clamorously into the presence of the Lioness, and demanded of her the settlement of the dispute. 'And you,' they said, 'how many sons have you at birth?' The Lioness laughed at them, and said: 'Why! I have only one; but that one is altogether a thorough-bred Lion.'*

*The value is in the worth, not in the number.*

Aesops. *Fables*

## Summary

The experiment tested two educational approaches to improving quality in clinical practice. The sum findings of the experiment are displayed below (figure 8a).

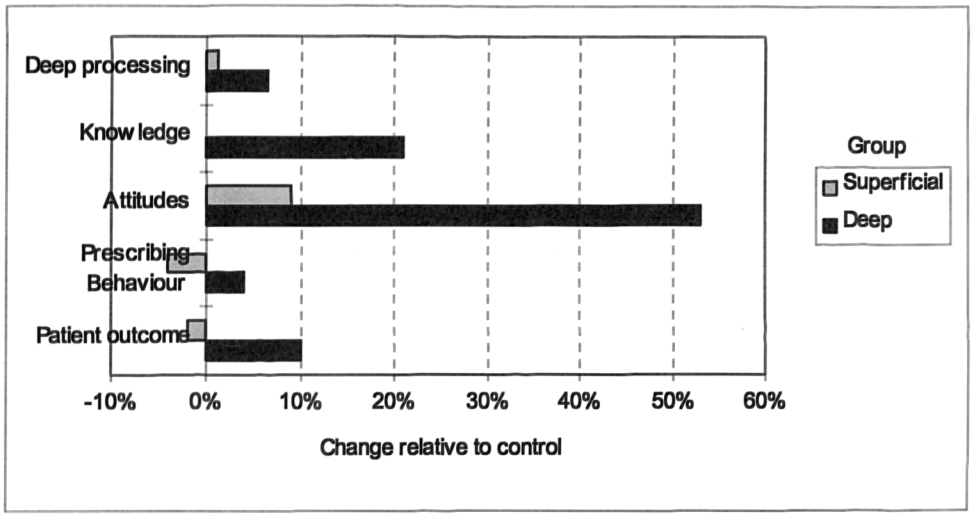



Figure 8a - Summarised results of change between pre- and post-intervention years in the Deep and Superficial groups relative to the Control (the median results of attitudes and prescribing are displayed).

There was a significant difference in depth of processing in the two intervention groups. Deep processing achieved greater knowledge of the guideline content, more favourable and persistent attitudes, consistent changes in behaviour, reduced prescribing cost, and improved patient outcome. Superficial processing did not enhance knowledge, did not produce generally favourable attitudes and behaviour, and did not improve patient outcome. There was evidence of some adverse effects from superficial processing, namely susceptibility to prescribing newer commercially-promoted antibiotics, perseverance in prescribing obsolete antibiotics, and increased prescription costs. An additional benefit of deep processing was the accumulation of criticisms of the guidelines drawn from clinical experience which

are being used to develop a second edition. In short, the extra mental effort required to deeply process information showed superior effects in every relevant parameter compared to superficial processing.

## 8.1 Interpretation of results

Of the three experimental groups the control was isolated from the interventions; the superficial group was engaged in a superficial level of processing of the guideline content and issues by means of an hour-long medical lecture event; and the deep group was engaged in a deeper level of processing of the guideline content and issues by means of hour-long medical visits.

All figures in this section (8.1) show favourable changes as positive. Consequently, where *decreases* are favourable (e.g. reduced prescribing of new antibiotics) the polarity of the Y axis is reversed. Similar graphs have identical scales and are shown together to aid comparison.

Deeper processing of the issues was achieved in the deep group than the superficial - more novel comments - and much more time was devoted to deep processing in that group.

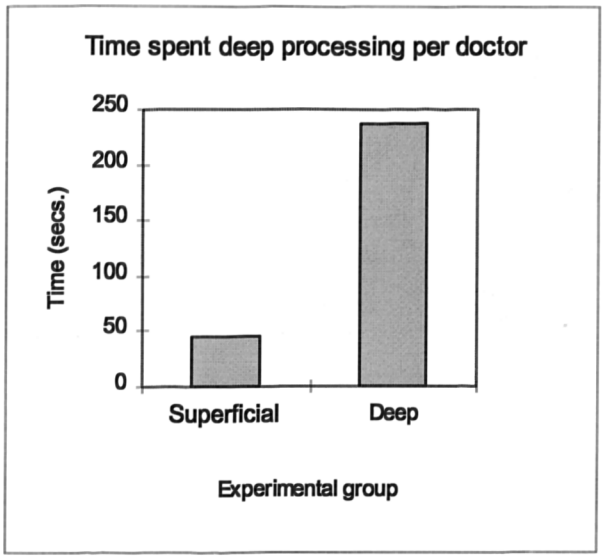



**Figure 8.1a Mean time (seconds) devoted to verbal expressions of deep processing per general practitioner in each group.**

Having achieved significant differences between the groups in the independent variable (deep processing) the experiment sought evidence of favourable and unfavourable changes in the conjectured dependent variables; knowledge, attitudes, prescribing and special investigation behaviour, and patient outcome.

### 8.1.1 Knowledge

There was a dramatic improvement in overall knowledge of locally perceived best practice in the management of infections. The superficial group showed no improvement post-intervention, whereas the deep performed **21%** better than expected.

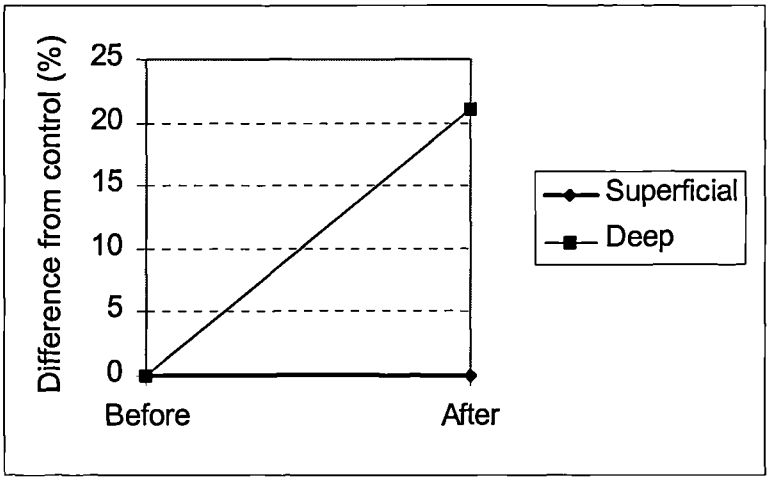



Figure 8.1.1a The performance of each intervention group relative to control on recalling best practice for common infections.

### 8.1.2 Attitudes

Changes in attitudes to the treatment of common infections were mixed in the superficial group, but universally favourable in the deep. Median attitude change was 9% more favourable in the superficial group relative to control, and 53% more favourable in the deep.

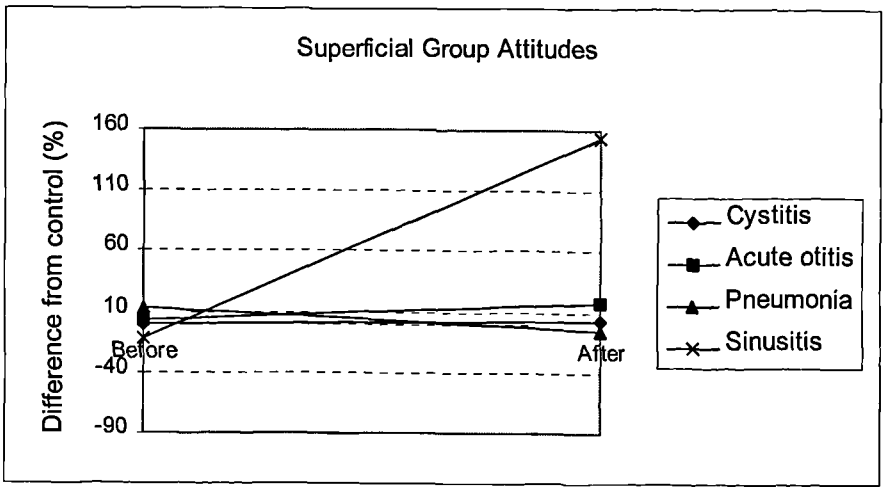



Figure 8.1.2a Superficial group attitudes relative to control for common infections.

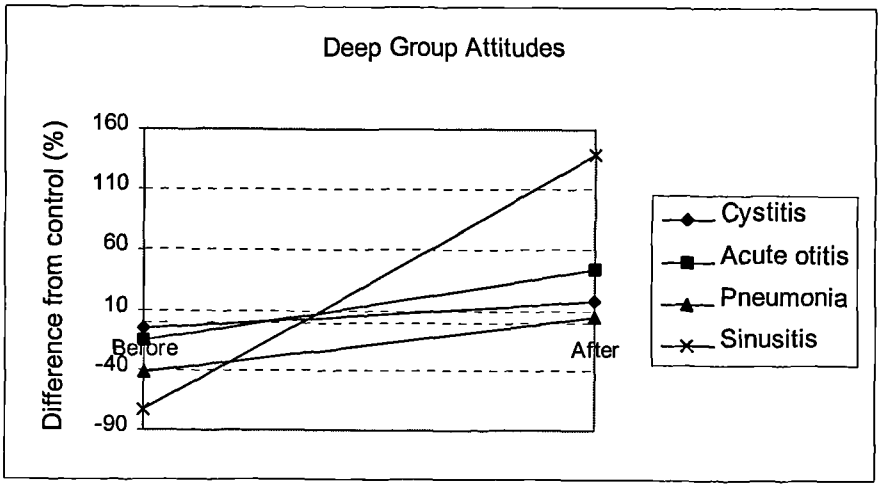



Figure 8.1.2b Deep group attitudes relative to control for common infections.

### 8.1.3 Prescribing behaviour

Superficial group prescribing performance in the 12 months post-intervention (relative to control) was *adverse* for 6 indicators and *favourable* for only 3 (co-amoxiclav, trimethoprim and new antibiotics). Conversely, deep group performance was *favourable* for 7 indicators and *adverse* for only 2 (amoxycillin and co-amoxiclav). The difference in prescribing change between the superficial and deep groups was also in the superficial group's favour for amoxycillin and co-amoxiclav. However, the attitude questionnaire identified that this may be due the fact that the superficial group had been prescribing these drugs more frequently for indications not promoted by the guidelines - more superficial than deep group respondents would use co-amoxiclav in acute otitis media, for example, whereas the guidelines recommended amoxycillin.

Median behaviour change was 0.9% less favourable in the superficial group and 0.6% more favourable in the deep than expected from control.

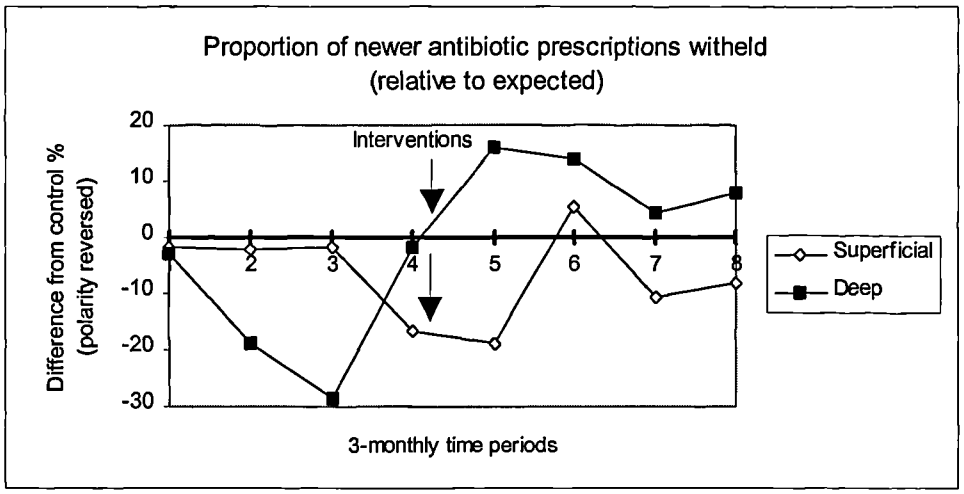



Table 8.1.3a Superficial and deep group prescribing performance (relative to control) for reduced prescribing of newer antibiotics per 3-month period.

Whereas superficial group performance on reducing new antibiotic prescribing seems favourable a breakdown of the results in 3-monthly time-periods shows the major change to have occurred a time remote from intervention. Therefore the intervention (superficial processing) and the effect (reduction in newer antibiotic prescribing) cannot be causally related.

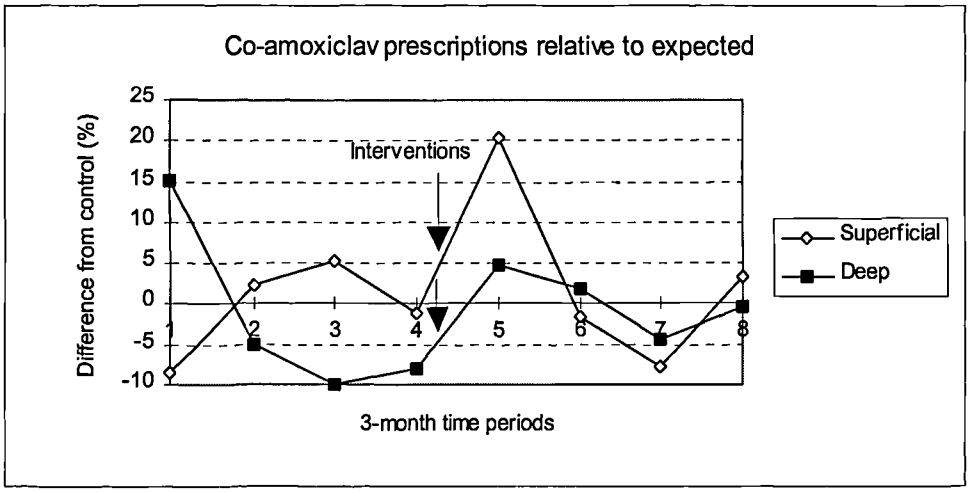



Table 8.1.3b Superficial and deep group prescribing performance (relative to control) for proportion of prescribing of co-amoxiclav per 3-month period.

The superficial group's good performance on the co-amoxiclav prescribing indicator is shown, by the same procedure, to be attributable to a large response in the first 3-month period which did not persist. A similar, but lesser, phenomenon is shown by the deep group.

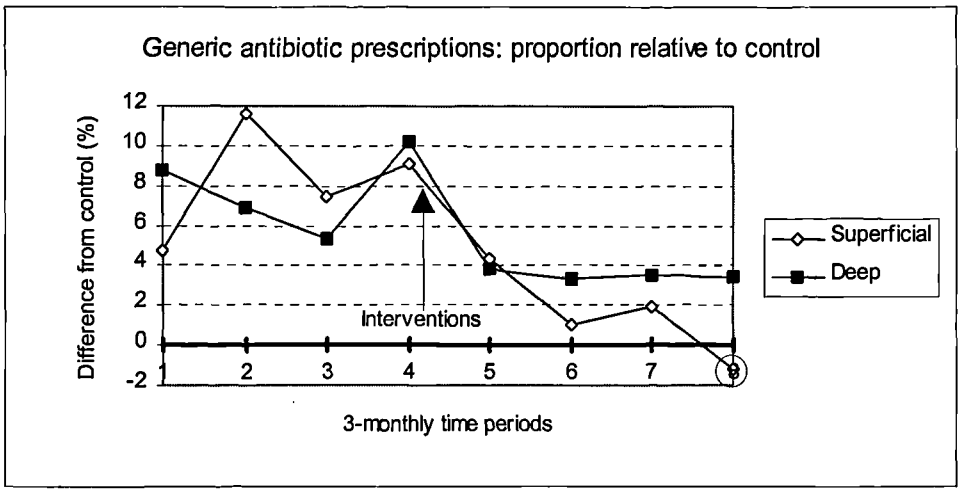



Table 8.1.3c Superficial and deep group prescribing performance (relative to control) for proportion of antibiotics prescribed generically per 3-month period.

Performance on generic prescribing changed adversely in the deep group in the 12 months post-intervention. However, examination at 3-monthly intervals shows that the deep group maintains a better than expected (from control) position, while the superficial group performs worse than control in the 9 to 12 month post intervention period (period 8).

### 8.1.4 Patterns of laboratory investigation

The superficial group displays minor (non-significant) changes in the proportion of requested microbiology investigations which might be affected by adherence to the guidelines. The deep group however, displays a significant favourable increase in gastro-intestinal investigations of 10%.

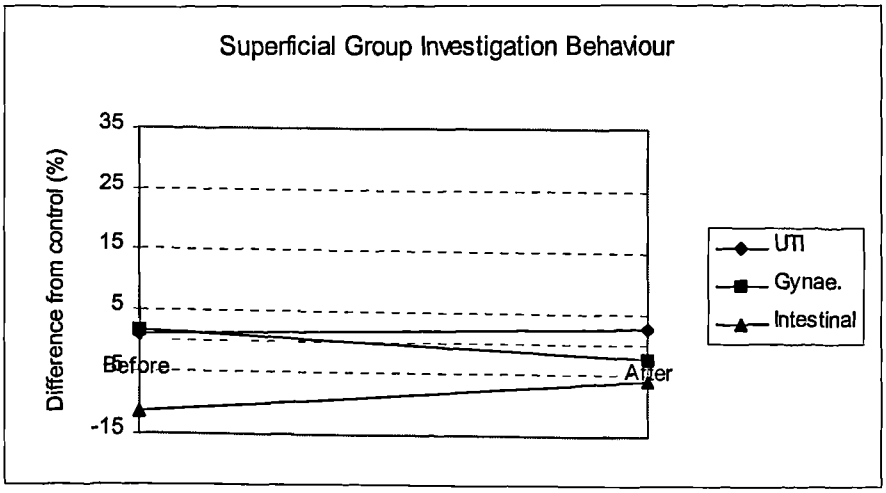



Table 8.1.4a Superficial group proportions of investigations consistent with major recommendations in the guidelines.

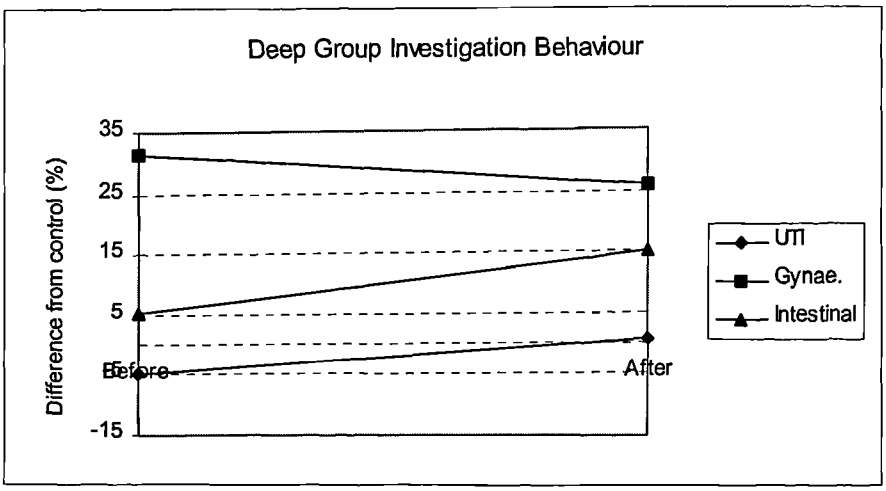



Table 8.1.4a Deep group proportions of investigations consistent with major recommendations in the guidelines.

The guidelines recommended referral of patients with pelvic infections for specialist investigation, however this recommendation was heavily criticised by general practitioners in the deep groups as impractical, so it is not surprising to see an adverse effect in the deep group. The guidelines recommended empirical antibiotic treatment of urinary tract infections, but did not

state whether a urine sample should be sent to the laboratory simultaneously or not. With this equivocal state of affairs, again it is not surprising to observe only a modest favourable response in the deep group.

## 8.1.5 Outcomes

Clinical study of a common indicator infection, cystitis, showed a clinically significant improvement in performance after deep processing.

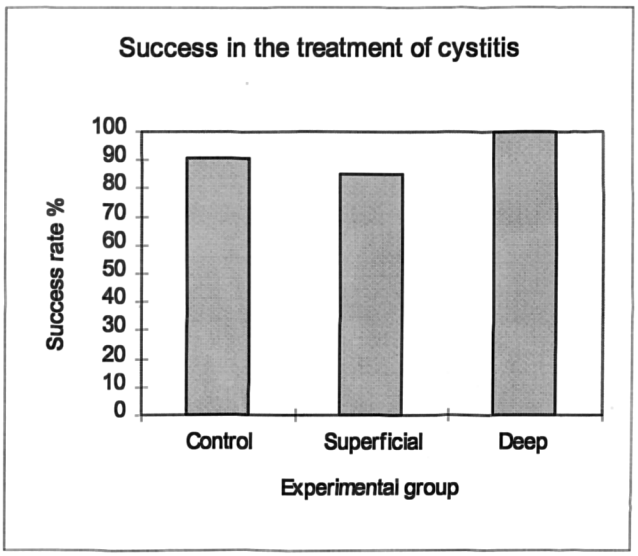



Figure 8.1.5a Success rate in the treatment of cystitis after one consultation.

It is apparent that none of the above changes after deep processing were associated with increased cost; on the contrary, deep processing resulted in a saving per antibiotic prescription. The cost per prescription increase in the deep group over the study year was 53% less than expected from control, whereas the superficial group increased 69% more.

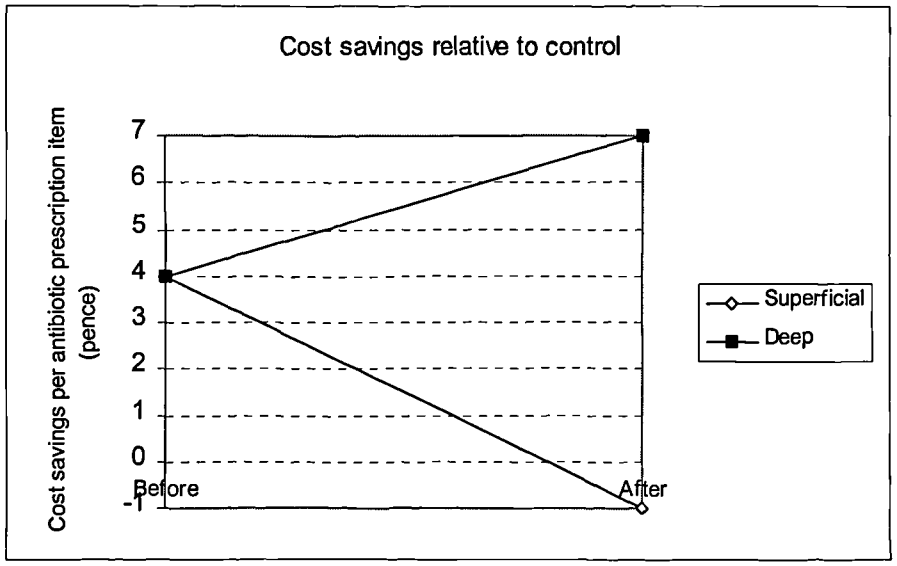



Table 8.1.5b Cost savings per antibiotic prescription (difference in cost per item relative to control with reversed polarity).

Both intervention groups issued 4p less expensive prescriptions than control before intervention, but only deep processing was associated with further savings after.

This saving per antibiotic prescription, if extrapolated, would represent savings of £12,000 to the district.

### 8.1.6 Summary of results (all endpoints)

It has been established that processing of the guideline messages was deeper in the deep group than the superficial (section7.1). Subsequently knowledge of the 'correct' management of common infections (as stated in the clinical guidelines) improved by 21% following deep processing, and not at all following superficial processing.

Median attitude regarding treatments for a range of infections was 41% more favourable following deep processing (range 23% to 212%), but only 2% more favourable following superficial (range -17% to 167%).

Median antibiotic prescribing behaviour was 5.3% more favourable following deep processing (range -7.5% to 5%), and 2% more favourable after superficial (range -17.2% to 7%).

Investigation behaviour of intestinal infection was 10% more favourable after deep processing, but not significantly improved after superficial.

Patient outcome was shown to be improved in an indicator infection (cystitis) by 10% after involvement of general practitioners in deep processing, but not at all after superficial (section 7.7.1).

Cost per antibiotic prescription increase in 1993 to 1994 (the study period) was less after deep processing:

| Group | Increase in cost per antibiotic prescription. |
|---|---|
| All England | 27p |
| Control group | 7p |
| Superficial group | 12p |
| Deep group | 4p |

Deep processing displayed superior effects in all groups of relevant variables, including subsequent clinical outcome. The changes in all indicators in the main study relative to changes in the control group are displayed in table 8.1.6a.

Table 8.1.6a Percentage changes in all indicators after the interventions: difference from control changes. Where reductions rather than increases were advocated the polarity has been reversed to aid analysis.

| *Data* | **Indicator** | **Superficial group** | **Deep group** | *Difference* |
|---|---|---|---|---|
| *Attitudes:* | Cystitis | 2 | 18 | *16* |
| | Acute otitis media | 8 | 30 | *22* |
| | Bronchopneumonia | -9 | 24 | *33* |
| | Sinusitis | 23 | 35 | *12* |
| *Prescribing:* | Amoxycillin | -2.4 | -0.7 | *-1.7* |
| | Co-amoxiclav | 0.6 | 0.2 | *-0.4* |
| | Trimethoprim | 0.1 | 0.6 | *0.5* |
| | Guideline antibiotics | -1.8 | -1.8 | *0* |
| | New antibiotics | 0.2 | 0.9 | *0.7* |
| | Co-trimoxazole | -1.4 | 3.1 | *4.5* |
| | Tetracyclines | -0.19 | 0.26 | *0.45* |
| | Generic antibiotics | -1.8 | 0.9 | *2.7* |
| | Cost saving | -1.1 | 1.6 | *2.7* |
| *Laboratory:* | Urine tests | 0.8 | 3.2 | *2.4* |
| | Gynaecological tests | 0.6 | 1.4 | *2* |
| | Intestinal tests | 0.4 | 0.8 | *0.4* |
| ***Mean*** | | **1.13** | **7.34** | *6.08* |

The descriptive statistics for all indicators are displayed in table 8.1.6b.

Table 8.1.6b Summary statistics for all indicators in the main study. Where reductions rather than increases were advocated the polarity was reversed before analysis.

| | Mean | Lower to upper 95% CL | Median | Minimum to maximum value |
|---|---|---|---|---|
| **Superficial** | 1.13% | -2.5 to 4.7 | 0.2% | -9 to 23 |
| **Deep** | 7.34% | 0.9 to 13.79 | 1.15% | -1.8 to 35 |

**Analysis of variance**

Where a study relies upon multiple endpoints, isolated examination of individual indicators may lead to a type 1 error where an indicator happens to have statistical significance, or a type 2 error where it does not. The study had 16 indicators; even if the null hypothesis were true there would still be a good chance of throwing up at least one significant difference. To address this issue the changes (in table 8.1.6a) were subjected to an analysis of variance (ANOVA).

The data is skewed (witness the difference in means and medians in table 8.1.6b) so a non-parametric test was most appropriate. The Kruskal-Wallis test, a one-way, non-parametric analysis of variance, showed that the deep group yielded significantly larger changes than the superficial group (mean difference = +6.08%, 95%CI = +1.93% to +10.23%, P = 0.014).

# Chapter 9

# Conclusion

*Though the world does not change with a change of paradigm, the scientist afterward works in a different world.*[209]

Thomas Kuhn, philosopher.1962

## Summary

Clinical guideline strategies, should incorporate involvement of the intended recipients in deep processing (effortful thinking) of guideline content at some, or every, stage. Superficial processing should only be used to make the recipient aware of issues and their relevance to him, hence to motivate deep processing.

## 9.1 Hypothesis tested

There was significantly more verbal evidence of deep processing in the deep group than in the superficial; the comments were predominantly novel whereas those in the superficial were confined to relevant questions, and more time was devoted per participating general practitioner in verbal contribution. Only the deep group recalled appropriate infection management better than control after intervention.

Twelve months post-intervention attitudes relative to control were statistically significantly more favourable in the deep group, excepting only sinusitis which was more favourable in the superficial (paradox I)*. *All* attitudes changed significantly in the deep group post-intervention, but *none* in the control or superficial.

Prescribing performance for the deep group was significantly better than control for all the guideline drugs, excepting co-amoxiclav. The same is true for the deep relative to the superficial group, except for co-amoxiclav again. However, guideline drugs

*Paradox I - after examination of the data in three month segments it is apparent that there was a dramatic rise in co-amoxiclav prescriptions in the first 3 months in the superficial processing group which returned to previous levels subsequently. This significant, but short-term effect is characteristic of successful superficial persuasion.

were adopted more often in superficial than deep (paradox II)*. Deep group prescribing of guideline drugs was significantly more favourable than control.

Further, new drugs were prescribed significantly *more* by both control and superficial groups, whereas the deep group prescribed significantly *less* of these drugs. This indicates an enhanced resistance to commercial marketing unique to the deep group.

Only the deep group prescribed significantly less tetracyclines and co-trimoxazole. This displays a rational abandonment of 'obsolete' drugs (superseded on side-effect profile by other antibiotics) unique to the deep group.

Prescribing data in the period 9 to 12 months post intervention analysed by class shows that all the favourable effects described above have persisted for over 9 months. In the 9 to 12 month post-intervention period the deep group preferentially prescribed significantly more of the recommended penicillins and

---

*Paradox II - Examination of the questionnaire data showed that although the superficial group prescribed a greater proportion of guideline antibiotics, they were often for the wrong conditions (eg co-amoxiclav for otitis media). This apparent paradox displays how caution must be exercised when employing collective indicators for prescribing (PACT) analysis.

cephalosporin - as proportions of those classes of drugs - than either control or superficial.

Mid-stream specimens of urine, vaginal swabs and faecal samples make up the majority of microbiological investigations sent by general practitioners. The proportions of these investigations changed appropriately in the superficial group, but most in the deep processing group as predicted. Quite startling was the large increase in volume in all three groups of 25%; this presumably reflects the increasing recognition of the value of conventional microbiology test results in general practice.[210] In spite of this element of statistical 'noise' the differences in changes in the three groups were significant enough to be detectable.

The definitive test of the effectiveness of a guideline is clinical benefit to the patient. Clinical audit of cystitis showed significantly less relapses after empirical treatment in the patients of deep group practitioners.

The general practitioners in the deep group made many useful comments on the content and format of the guidelines. The comments would not have been apparent in the main from appraisal solely of published scientific literature. The weakness of pure scientific 'evidence-based medicine' was thus exposed, and

the value of clinical acumen in the development of clinical guidelines demonstrated. Of particular interest was the input of the practice nurses involved in deep processing. The nurses' comments covered entirely different and equally important aspects of infection management to the doctors and displayed their contribution to general practice. Several commented upon how useful they had found the guidelines and this perhaps reflects a general enthusiasm amongst nurses for clinical guidelines.[211]

Only the deep group prescribing cost was less than control per head of population (by 1.5%). Cost per item changes were also favourable in the deep, but not the superficial, group. The deep processing intervention was more cost effective than the superficial.

In summary; relative to the control group deep processing was associated with favourable attitudes toward the issues. These attitude differences had persisted for over 12 months. There were no significantly different favourable attitudes in the superficial group.

The favourable attitudes in the deep group were associated with consistent prescribing behaviour. Furthermore, there was evidence of resistance to counter-persuasion as the deep group

were the only group to significantly reduce prescribing of newer (and most heavily commercially promoted) antibiotics.

Table 9.1a The characteristics of the deep processing group (relative to control) not evident in the superficial.

| **Consequences of deep processing:** |
|---|
| • Persistent memory of facts (knowledge). |
| • Persistent favourable attitudes. |
| • Consistency of attitude and behaviour. |
| • Improved patient outcome. |
| • Resistance to commercial drug promotion. |
| • Abandonment of obsolete practices. |
| • Lower cost per prescription item. |
| • Beneficial criticism of guidelines. |

A review article by Conroy and Shannon states that any strategy for implementing guidelines should have an impact at four levels:[212]

- Increasing knowledge, that is , making clinicians aware of the guidelines
- Changing attitudes, such that clinicians agree with and accept the recommendations as a better standard of care
- Changing behaviour, such that clinicians change their clinical practice to conform with the guidelines
- Changing outcomes, by improving patient health and quality of care (sic).

It can be seen from table 9.1a that the impact of a single strategy, deep processing, will achieve these and more. Much of the literature on successful implementation concentrates on procedures which are means to deep processing, such as

'involvement' of the recipients in development or implementation. However, these activities have been elevated to the status of goals in themselves.[213] Exploring new ways of achieving deep processing may open up opportunities for effective change management in general practice.

The findings were predicted by the Elaboration Likelihood Model of persuasion which holds that deep processing of strong messages results in attitudes which are persistent, resistant to counter-persuasion attempts, and are followed by consistent behaviour.[214]

The sequence of events described above corroborates the hypothesis that deep processing results in better recall of information; persistent and resistant attitudes; and consistent behaviour. The extensive hypothesis testing by attempted refutation in chapter 7 confirms this. Furthermore, patient outcomes can be improved without incurring increased treatment costs.

## 9.2 Achievement of aims and objectives

Critical to success in a scientific experiment is control. Relevant variables must be balanced across all groups or samples as well as possible apart from the variable under study which must be present in the intervention group exclusively. Given an infinite number of subjects a random process will allocate variables increasingly equally. The advantage of randomisation is that it distributes variables, both known and unknown, equally. Even though all practices (bar one exclusion) were empanelled in the experiment the numbers in each group were insufficient to rely upon pure randomisation. Selection of practices according to variables possessed achieves a quantifiable distribution, but of known variables only. The experiment adopted a compromise method relying mainly on randomisation, but with a few reassignments to balance the known confounding variables. The effect was the best distribution possible. The final random allocation of the three group labels (control, superficial and deep) prevented bias. The practice rather than the practitioner was chosen as the data unit for this study, in common with other studies of general practitioner behaviour it was felt that data is unreliable below practice level.[215]

For antibiotic prescribing behaviour the chosen variable was 'items' or individual prescriptions. 'Items' have short-comings because of variations in lengths of courses and for many drugs the variation is significant. However, comparison of quantity of tablets per item for antibiotics across all Family Health Services Authorities shows a modest variation (median coefficient of variation = 5%, range 3% to 8%) compared to all categories of drugs (median coefficient of variation = 8%, range 3% to 26%).[216] In any case the primary indicator of change was the number of times a named drug was prescribed; the quantity per prescription (length of course) was secondary. For these reasons 'items' were felt to be the most appropriate measure of change.

Other potential sources of bias are the availability of data to the researchers during the intervention and unconscious application of greater effort into the favoured group. The former was successfully protected against by delaying data analysis until the end of the experiment, the latter by having separate educational teams apply the two interventions.

Time exposed for each subject was kept strictly to one hour, whether lecture or interview. The educational messages were identical for both groups. Deep processing was allowed in both groups, however this was the variable under test and was

successfully generated to a significantly higher degree in the deep group.

To be applicable in practice the experimental settings were as realistic as possible. The lecture and practice visit formats were familiar to general practitioners, although the presence of television cameras may have caused some uneasiness at first.

The attitudes and behaviour of the three groups approximated reasonably well pre-intervention (as can be seen from table 7.1 in Chapter 7). However, that some significant differences already existed prior to intervention makes interpretation of the results more difficult. To this end the data is analysed three ways in an explicit manner; between-groups before, between-groups after, and change within-groups.

The Family Health Services Authority's register of patients per practitioner proved to be unreliable. The Government Statistical Service[217] and the Audit Commission[218] have all prefaced reports on prescribing with warnings concerning this problem, which has been described in depth by the national Prescribing Research Unit.[219] A solution is to avoid data presentations and analyses which depend upon a population denominator wherever possible. This was achieved by employing Fisher's exact and Chi squared tests to aggregated group results.

## 9.3 Significance of the experiment

### Significance for clinical guidelines:

The experiment explores two areas where it is recognised there are few established facts. Firstly, little is known about the effect or validity of locally developed guidelines because current evidence is conflicting; and secondly, the best method of implementing guidelines has not been satisfactorily established.[220] Further, although the significance of 'involvement' is clearly recognised (it is a prominent feature in the summary table in the Effective Health Care review of guidelines)[221] the effect of involvement is variable, suggesting that it is the *nature* of the involvement is important. This thesis hypothesised that *involvement of recipients in deep processing of the issues is essential for success in clinical guideline implementation.*

Deep processing is the engagement of individuals in effortful thinking on the strength and personal relevance of messages. Reviewing previous guideline research does not shed much light on this aspect as often relevant details have been neither sought nor recorded.

Of the 59 'rigorously evaluated' guidelines in Grimshaw and Russell's review article, 48 (81.4%) evaluated effect on the process, and 11 (18.6%) on outcome[222]. Where process has been evaluated, the criterion for success has been compliance with the guideline as measured directly[223] or by proxy through changes in prescribing or referral rates.[224] The implicit assumption is that the guidelines themselves do not need to be evaluated. Where the guidelines have arisen from 'expert' concensus delivered to other doctors working in that field this could seem arrogant. Where guidelines have arisen from published scientific evidence this assumes that the controlled environment of the experiment will also be experienced in practice. The assumption that users of guidelines have nothing to contribute fails to capitalise on a wealth of experience and clinical acumen. To reinforce this loftiness, authors have often blamed non-compliance on poor education of recipients[225] [226] and have not questioned whether the guidance was actually suitable for application in practice.

Where outcomes have been evaluated, guidelines have not been compared with other guidelines (for the same medical conditions), but only against no-guideline controls. To discover that 'something is better than nothing' is to have confirmed that which is common sense. Where guidelines are altered it is in

response to shifts in expert opinion, or increasing scientific knowledge[227], but not any reported problems with applying the guidance in practice.

Guidelines are a means to an end, namely to achieve the most effective care for patients. This end is the aspiration of all doctors. How effective guidelines are as a means to achieve that end is seldom inquired of guideline recipients by authors. An opportunity to spot defects in guidelines and refine them is ignored. In short, guidelines have been extensively evaluated with regard to compliance but not utility.

It appears that guidelines are here to stay and, as these are an aid to practice, should be a good thing for doctors. However, the predominance of economic evaluations in Government-funded guidelines is a potential threat to doctors. Health economists may develop guidelines in good faith based upon published scientific evidence, assumption-riddled models, and the 'rational man' safely remote from actual illness and distress, but adherence to such guidelines is likely to achieve sub-optimal benefits in real consultations. It is imperative that practising doctors have control over the guideline process as established in this experimental design. However, a balance must be struck between scientific

evidence and practical experience; the experiment explores the process of synthesising science and practice.

The educational impact of superficial processing of guideline issues relative to deep has not been explored before. The didactic lecture is assumed to be ineffectual relative to interactive work, but rarely have the two been directly compared. Particularly, economic comparisons between the two have not been attempted before. The new features of this study, not evident in previous published work on guidelines, are shown in table 9.2a.

Table 9.3a Distinctive features of this study.

| **Unique features of the experiment** |
|---|
| • The principal variable was the subjects' *depth of processing*. |
| • Evidence of *persistent* and *resistant* changes in both attitudes and behaviour were sought 12 months post-intervention. |
| • The intervention encompassed an *entire clinical field* in medicine. |
| • The validity of the guidelines was *tested by recipients* as well as experts. |
| • Two realistic educational options for implementing the guidelines were compared with a control. |
| • An *eclectic* approach to design and analysis was undertaken, incorporating methods and theory from several academic disciplines. |
| • Feedback and repetition were avoided. |

## Wider significance of the study: a paradigm shift in medical education

This thesis presents a unifying principle underlying the often disconnected, but concerted belief of many current medical

*educators that doctors learn best by tackling problems, not by* being fed information. Whether action learning, problem based learning, detailing visits, clinical discussions, clinical audit review, small group work, project work, case study conundrums, is being advocated, the common thread is deep processing. The experiment is consistent with the current trends in medical education towards a deep processing and community-centred approach.[228] These new approaches represent a new paradigm in medical education. The innovators have become dissatisfied with the failings of the old paradigm; that of the lecture, the bulletin, the lecture notes, the books of lists, the pronouncing tutor, the role model, the expert authority, repetition - the paradigm of superficial processing.

However, as Thomas Kuhn wrote - To reject one paradigm without simultaneously substituting another is to reject science itself (sic).[229] The pioneers are exploring new ways out of frustration at the failings of the old, but most will not completely abandon the obsolete until the new paradigm, currently a collection of new and effective methods, is underpinned by a rational and robust theoretical framework. Deep processing, as tested in this experiment, provides that.

I, with research colleagues, have previously shown that deep processing improves memory for medical facts[230], and it changes clinical behaviour in a prolonged manner.[231] In the main experiment of this thesis the old paradigm is pitted against the new in an arena of controlled scientific experiment, and their fundamental components, superficial and deep processing, are polarised by the interventions. In many areas where an educator would expect an effect, such as knowledge, attitudes, and clinical outcome, the superficial paradigm failed to produce any effect. The deep paradigm, however, delivered favourable effects in every area and where both paradigms were successful, the deep was usually superior.

Much work is still required to delve into deep processing further and to expand the theory system to support it. However, perhaps the message to us all as educators is that we should concentrate more on selecting and testing *differing* ways to achieve deep processing when we undertake *differing* educational projects, rather than applying our preferred educational deep processing method to *all* our work. The various interactive methods are merely *means* to deep processing.

Kuhn warns that when a new paradigm becomes widely accepted, scientists often reject the old completely. This would be

a mistake. Superficial processing has a place in medical education; namely to raise awareness of new issues and words *prior* to deep processing.

If this thesis helps the paradigm shift to occur in some small way, that would be a most satisfactory outcome.

## 9.4 Implications

The requirements for the necessary features of valid guideline development have been described and include systematic evaluation of all relevant literature, ranking of options and economic analysis. As time goes by the requirements are becoming so exacting that they far exceed the scientific competence currently expected of any individual general practitioner. The widening gulf in performance between those who adopt 'perfect' guidelines (they will require computers to do this) and those who do not will give rise to some interesting professional and legal problems. Clinical guidelines and increasing expectations will conspire to automate medicine; the profession stands at the brink of a new electronic age of medicine, appropriately at the turn of the Century. This will constitute a 'paradigm shift' in medical information.[232] Kuhn predicts that those dissatisfied with current (relatively informal) consensus medicine will enthusiastically leap to embrace the new formal rigorous guideline science. Equally those comfortable with the established paradigm will resist the new paradigm vigorously. Popular articles on guidelines display where the battle-lines are drawn; between those who see rigorously developed evidence-based guidelines as the professions' salvation,[233] and those who see them as a threat to clinical freedom and a potential source of perpetual torment.[234]

The conflict has been summarised by Gene Feder who urges caution.[235] Who is right is immaterial, history teaches that technological progress can be delayed sometimes, but prevented never. Faced with the inevitable, the profession is best advised at least to ensure that guidelines are driven by patient benefit and not financial efficiency, although often the result may satisfy both. There is a movement amongst health service managers to reinforce guidelines by introducing them into contracts with providers.[236]

Charlton points out the folly of absolute faith in medical research findings and recommendations, and equally the problems with consensus methods of establishing 'evidence' (subjective opinion), no matter how elegantly designed.[237] The difference in the approach to guideline development described in this thesis is the prominence of clinical acumen. Scientific evidence, where available, dominates the guidelines, but is viewed in the light of collective and individual clinical opinion. Furthermore, the guidelines do not assume validity because they are scientific 'evidence-based', but are tested in application. (Guideline review) criteria can be used to aid implementation of guidelines by providing a standard against which to monitor performance and enabling clinical audit (sic).[238] Regardless of the

weight of science behind a guideline, the essential measure of validity is whether patients benefit from it. Dans mentions the description of guidelines as 'cookbook medicine' and follows with 'the proof of the pudding is in the eating' - ergo the recent emphasis on specifying and validly measuring outcomes (sic).[239]

The implications for guideline development are listed in table 9.4a.

Table 9.4a Implications for clinical guideline development.

| **Implications for guideline development:** |
| --- |
| • Guidelines must incorporate science and clinical experience.<br>• Guidelines should be tested by intended users in practice.<br>• Guidelines must be perpetually refined and updated according to accumulating scientific evidence and clinical experience.<br>• Guidelines must be explicit in order to invite criticism. |

Of course a guideline can only be effective if it is implemented. As discussed above, guidelines may ultimately operate on electronic patient management systems. The doctor could automatically follow the machine's default settings unless his clinical acumen or his patient's unique requirements dictate another option. To safely and appropriately opt out he must be aware of the rationale behind the guidelines programmed into the machine, otherwise how will he know when the default guideline is inappropriate? In an electronic future guidelines will need to be

associated with medical education which promotes understanding through deep processing. Presently deep processing is vitally important; in the absence of mechanical aids the doctor must not only *remember* and *understand* the guideline issues, but must be persuaded to actually *act* on them. Fortunately deep processing is known to achieve all three objectives. Many reviews of guidelines talk vaguely of 'involvement', this thesis clarifies the nature of the involvement required for the first time (i.e. involvement in deep processing). An added benefit for the guideline editor is that achieving deep processing amongst his guideline users unlocks a wealth of clinical acumen that can criticise, refine, and make the guideline more practicable. The integration of individual recipients, alone or in small groups, has the advantage of being a powerful educational activity; where deep processing is employed to solve important problems.[240] The implications for the implementation of clinical guidelines are listed in table 9.4b.

Table 9.4b Implications for clinical guideline implementation.

| Implementation: |
| --- |
| • Users must be involved in deep processing of the guideline content at some stage.<br>• Simply publishing guidelines is insufficient.<br>• Didactic transmission of guideline messages has a limited and short-lived impact on both attitudes and behaviour.<br>• Deep processing of guidelines has a significant, persistent and resistant impact on both attitudes and behaviour. |

## 9.5 Limitations and further work

There is no such thing as a perfect guideline because scientific evidence is never perfect, differences of opinion will always exist. Even ranking of evidence according to the type of experimental design does not guarantee choice of the correct option, it merely preferentially promotes evidence arising from double-blind randomised controlled trials whether correct or not. There is always a point where the options open have to be reduced by subjective preference or application of principles. For example one guideline might reasonably promote clarithromycin over metronidazole as the antibiotic of choice in helicobacter eradication therapy on the grounds of greater likelihood of success in treatment of individual patients (principle A).[241] Another may promote metronidazole over clarithromycin on the grounds of restricting the use of newer antibiotics in general practice to prevent bacterial resistance problems, therefore protecting the general public interest (principle B).[242] A majority of patients may benefit from rigid guidelines at the expense of a minority of unfortunate incongruous patients; an ethical blunder known as the 'tyranny of the majority'.[243] These dilemmas illustrate the ethical implications of clinical guidelines, an unexplored aspect.

Another dilemma faced in this study is realism versus idealism. To be applicable elsewhere the interventions had to be realistic and familiar. However, this could have introduced a bias. Thirty-three percent turnout at a lecture is as high as one could reasonably expect; however, 66% were not exposed to the educational intervention. By contrast less than 10% of the deep group were absent when visited. Examination of the subset of the superficial group who had attended the lecture and completed both attitude questionnaires demonstrated no more favourable changes in attitude than in the superficial group as a whole. It is therefore unlikely that the lecture had any lasting effect on attitudes (the change was measured at 12 months) and therefore could not have contributed to any changes in behaviour observed in the superficial group except perhaps those of short duration. Also, from a pragmatic point of view, if high attendance is only guaranteed by visiting then lectures should probably be dropped in favour of visiting where feasible anyway.

The guidelines were only tested on Wirral. Their success could have been due to either involvement of the practitioners in their development, or implementation, or both. Implementation of the next guidelines edition in another district would test their universality. Equally only one clinical field was tested; the

effectiveness of the method in another field, such as cardiovascular disease or dermatology, should be studied to see if the process is truly generalisable to the rest of medicine.

The computer age is rendering printed information obsolete. Further work is required to install the 'Microbes' guidelines into a patient management system to see if they work. After development simple provision of the electronic guidelines should be compared with provision plus deep processing. Perhaps interactive programmes can facilitate deep processing while still automating routine decisions.

The study applied the same educational techniques to all practitioners in a group. People have different preferred learning styles and readiness to learn.[244] The effect of the lecture may have been dramatic on a couple of attendees, but swamped by the behaviour of the rest. Similarly deep processing may have adversely affected a few deep group practitioners, but again lost in the effect on the rest. More work is required to tailor education to individual practitioners' requirements.

Little is known of doctors' goals in life. Further work is required to explore this aspect of motivation. Martin Ford has developed a manual to enable the investigation of goals.[245] Alternatively, the approach adopted by modern sales executives

could be adopted. They do not assume their potential client's goals, instead they discover them during interview and emphasise the appropriate features of their product's profile to suit. Again this would be an interesting and useful avenue to explore, particularly for medical advisers and others who visit practitioners.

More work is remains to test patient outcomes and guidelines, even to the point of testing guidelines against each other. Unfortunately, unless a rapid method of appraisal is developed the guidelines are likely to be overtaken by new evidence before they are fully evaluated. The race to test guidelines scientifically before obsolescence is likely to be frantic until fully computerised and linked patient management systems allow rapid automatic audit (several years hence).

## 9.6 Lastly

Clinical guidelines can be made to work across an entire medical field. Scientific justification is essential and should be based on a review of the relevant literature. Recipient practitioners can identify impracticalities in the guidelines and test them in practice.

Involvement of the recipients in deep processing of the guideline content (through discussion in practice visits) results in favourable attitudes toward the content which persist for over 12 months, are resistant to pharmaceutical company counter-persuasion, and are followed by consistent favourable changes in clinical behaviour. The process is cost-effective and significant general improvements in clinical quality are apparent.

By comparison, superficial processing (lectures) has negligible long-term effect on knowledge, attitude, behaviour or patient outcome, and is not cost-effective.

To be effective clinical guidelines should involve recipients in deep processing during development, dissemination, implementation or audit. Involvement at all stages is probably not necessary. Systematic appraisal of relevant scientific evidence by academics and clinical practitioners should form the foundation of

guidelines. The value of practitioners' experience and acumen must be recognised and incorporated into the development or review process if guidelines are to achieve validity in practice.

Reliance upon mere publication of guidelines to change practice is an error, no matter how authoritative the source. This is also true of reliance upon printed bulletins, didactic lectures and the like which have all been shown to be ineffective except in the very short-term. Superficial processing alone is virtually valueless.

Deep processing is the key educational approach to improving clinical practice. The deeper the processing the more that messages are memorised, and form appropriate attitudes followed by consistent actions.

Having established that deep processing is the key ingredient it may be tempting to abandon superficial approaches altogether. Indeed the recent Kings Fund report on medical education[246] sees no place for the didactic lecture. However, deep processing does not arise spontaneously. Firstly a superficial appraisal of the issue is used by subjects to identify whether a topic is relevant or interesting, if so deep processing will follow. Successful deep processing strategies will initially exploit superficial processing to motivate the subject to engage in more effortful thinking.

There is growing dissatisfaction with the old paradigm of superficial medical education. However, while innovators in medical education have explored new techniques outside that paradigm and have had successes, a wholesale shift is unlikely until a new paradigm has emerged which has proved superior to the old. This thesis proposed that deep processing should be the new paradigm and has tested it scientifically. Deep processing proved not only to be effective, but to be significantly more effective than superficial processing in every way. More research is required to explore deep processing further and to develop and refine its theoretical base.

I commend deep processing to you, O reader. Beware however, once a person has shifted from the old paradigm to the new it becomes impossible for him to see things the same way again.

# Index

# References

All references are expressed in full (Vancouver style), in the sequence that they appear in the text, and duplicated where they are referred to on more than one occasion.

[1]Leibniz GW, cited in Ed Parkinson GHR. Philosophical writings (1973) p183 Everyman.

[2] Smith R. Audit and research. BMJ 1992;305:905-6.

[3] Harris C. Prescribing - a suitable case for treatment. Occasional Paper 24. 1985. Royal College of General Practitioners. London.

[4] Harris CM, Fry JS, Jarman B, Woodman E. Prescribing - a suitable case for prolonged treatment. JRCGP, 1985;284-7.

[5]Taylor RJ. Change in the established prescribing habits of general practitioners: an analysis of initial prescriptions in general practice. *British Journal of General Practice.* 1991;**41**:244-48.

[6]Waine C Ed. How to produce a practice formulary. (1989) The Royal College of General Practitioners. London.

[7]Gilleghan JD. Prescribing in general practice. Occasional Paper 54. (1991) p14.The Royal College of Practitioners.London.

[8]Martin AJ, Lacey RW. A blind comparison of the efficacy and incidence of unwanted effects of trimethoprim and co-trimoxazole in the treatment of acute infection of the urinary tract in general practice. *Br J Clin Pract* 1983;**37**:105-112.

[9]Griffiths LR, Bartzokas CA, Hampson JP, Ghose AR. Antibiotic costs and prescribing patterns in a recently commissioned Liverpool

teaching hospital. Part1: antimicrobial therapy. *Journal of Hospital Infection.* 1986; **8**: 159-167.

[10]Acheson HKW,Henley MH. Clinical knowledge and education for general practice. Occasional Paper 27. (1984) p6. The Royal College of General Practitioners. London.

[11]Thomas CGA. Medical Microbiology. (1973) p195. Bailliere Tindall.London.

[12] Rethans J-J, Saebu L. Do general practitioners act consistently in real practice when they meet the same patient twice ? Examination of intradoctor variation using standardised (simulated) patients. BMJ, 1997;314:1170-3.

[13]Kelley HH. Attribution theory in social psychology (1967) paraphrased in Zimbardo PG, Leippe MR. The psychology of attitude change and social influence (1991) p89-90 McGraw Hill. New York.

[14] Zimbardo PG, Leippe MR. The psychology of attitude change and social influence.1991. McGraw Hill.

[15]Forster DP, Frost CEB. Use of regression analysis to explain the variation in prescribing rates and costs between family practitioner committees. *British Journal of General practice.* 1991; **41**:67-71.

[16]Soumerai SB, Avorn J. Predictors of physician prescribing change in an educational experiment to improve medication use. *Med Care* 1987; **25**:210-21.

---

[17]Taylor RJ. (1981) MD thesis. University of Aberdeen. paraphrased in Gilleghan JD. Prescribing in general practice Occasional Paper 54. (1991) p1. The Royal College of General Practitioners. London.

[18] Taylor RJ. Prescribing costs and patterns of prescribing in general practice. Journal of the Royal College of General Practitioners. 1978;28:531-535.

[19] Weiss MC, Fitzpatrick R, Scott DK, Goldacre MJ. Pressures on the general practitioner and decisions to prescribe. Family Practice, 1996;13(5):432-8.

[20]Pitts J, Vincent S. What influences doctors' prescribing? Sore throats revisited. *Journal of the Royal College of General Practitioners.* 1989;**39**:65.

[21]Dickson D, Saunders C, Stringer M. Rewarding People:The skill of responding positively.1993. Routledge. London.

[22] Hill-Smith I. Sharing resources to create a district drug formulary: a countrywide controlled trial. British Journal of General Practice. 1996;46(406):271-5.

[23] Martin BC, McMillan JA. The impact of implementing a more restrictive prescription limit on Medicaid recipients. Effects on cost, therapy, and out-of-pocket expenditures. Medical care, 1996;34(7):686-701.

---

[24] Weiss MC, Fitzpatrick R, Scott DK, Goldacre MJ. Pressures on the general practitioner and decisions to prescribe. Family Practice, 1996;13(5):432-8.

[25]Mills G, Chaffe A. Is cost-awareness really improving? *Health Trends.* 1993;**25;1**:38-40.

[26] Gomez Perez B, Trilla Garcia A, Vernet Arias E, Corominas Garcia N et al. [Do physicians know the cost of the treatments and tests they prescribe?]. [Spanish]. Revista Clinica Espanola, 1996;196(8):523-8.

[27]Ford M. Motivation: Goals, emotions and personal agency beliefs. (1992)

[28]Bradley CP. Uncomfortable prescribing decisions: a critical incident study. (1992) *BMJ* **304**:294-6.

[29] Wilson RPH, Buchan I, Walley T. Alterations in prescribing by general practitioner fundholders: an observational study. *BMJ.*1995;**311**:1347-50.

[30] Wilson RPH, Hatcher J, Barton S, Walley T. Influences of prescribing characteristics on prescribing in fundholding and non-fundholding practices: an observational study. BMJ, 1996;313:595-9.

[31] Harris CM, Scrivener G. Fundholders' prescribing costs: the first five years. BMJ, 1996;313:1531-4.

[32] Surender R, Bradlow J, Coulter A, Doll H, Stewart Brown S. Prospective study of trends in referral patterns in fundholding

and non-fundholding practicesin the Oxford region, 1990 - 4. *BMJ.* 1995;**311**:1205-8.

[33] Stewart-Brown S, Gillam S, Jewell T. The problems of fundholding. Some benefits to patients but no effect on how doctors practise. BMJ, 1996;312:1311-2.

[34] David TJ, Patel L. Adult learning theory, problem based learning and paediatrics.

*Archives of Disease in Childhood.* 1995;**73**:357-363.

[35] Pedlar M, Boutall J. Action learning for change. 1992. NHS Training Directorate.

[36] Curry L, Putnam RW. Continuing medical education in Maritime Canada: the methods physicians use, would prefer, and find most effective. CMA Journal. 1981;124:563-6.

[37]Branthwaite A et al. Continuing Education for General Practitioners. Occasional Paper 38. (1988) p17 The Royal College of General Practitioners. London.

[38]Stanley I, Al-Sheri, Thomas P. Continuing education for general practice. *British Journal of General Practice.* 1993;**43**:210-214.

[39].Sibley JC, Sacket DL, Neufeld V, Gerrard B, Rudnick KV, Fraser W. A randomized trial of continuing medical education. *The New England Journal of Medicine.* 1982;**306**:511-5.

---

[40]Berbatis CG et al. Impact of a drug bulletin on prescribing oral analgesics in a teaching hospital. *Am J Hosp Pharm* 1982;**38**:98-100.

[41] Bloor K, Freemantle N. Lessons from international experience in controlling pharmaceutical expenditure. II: Influencing doctors. BMJ, 1996;312:1525-7.

[42] Braybrook S, Walker R. Influencing prescribing in primary care: a comparison of two different prescribing feedback methods. Journal of Clinical Pharmacy and Therapeutics, 1996;21:247-54.

[43] Santoso B. Small group intervention versus formal seminar for improving appropriate drug use. Social Science and Medicine, 1996;42(8):1163-8.

[44] Wensing M, Grol R. Single and combined strategies for implementing changes in primary care: a literature review. International Journal for Quality in Health Care, 1994;6(2):115-32.

[45] Davis DA, Thompson MA, Oxman AD, Haynes B. Changing physician performance: a systematic review of the effect of continuing medical education strategies. JAMA, 1995;274(9):700-705.

[46]Johnson MW. Mitch WE, Heller AH, Spector R. The impact of an educational program on gentamycin use in a teaching hospital. *Am J Med* 1992;**73**:9-14.

---

[47] Donner RS, Bickley H. Problem-based learning in American medical education: an overview. Bulletin of the American Library Association. 1993;81(3):294-8.

[48]Parsell G, Bligh J. Contract learning, clinical learning and clinicians. Postgraduate Medical Journal, 1996;72(847):284-9.

[49] Towle A, Cottrell D. Selfdirected learning. Archives of Disease in Childhood, 1996;74(4):357-9.

[50] Irvine D, Irvine S (Eds.). Making sense of audit. 1991. Radcliffe Medical Press. Oxford.

[51] Pedlar M, Boutall J. Action learning for change. 1992. National Health Service Training Directorate. Bristol.

[52] Newble DI, Entwistle NJ. Learning styles and approaches: implications for medical education. Medical Education, 1986;20:162-175.

[53]Craik FIM, Lockhart RS. Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behaviour.* 1972;11:671-684.

[54]Parkin AJ. Levels of processing, context, and facilitation of pronunciation. (1984) Acta psychologica 55:19-29 quoted by Matlin MW. Cognition (1989) p78 Holt, Reinhart and Winston Inc. Chicago.

[55]Onion CWR, Slade PD. Depth of information processing and memory for medical facts. *Medical Teacher.* 1995;17(3):307-313.

---

[56] Atkinson RC, Shiffrin RM (1968). Human memory: A proposed system and its control processes. In Spence KW, Spence JT (Eds). The psychology of learning and motivation: Advances in research and theory. **2**. Academic Press. New York.

[57] Soumerai SB, Avorn J. Predictors of physician prescribing change in an educational experiment to improve medication use. *Med. Care.* 1987;**25:**210-21.

[58]Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. In Bekowicz L (Ed.) *Advances in experimental psychology.* 1986;**19**:123-205.

[59]Zimbardo PG, Leippe MR. The psychology of attitude change and social influence. (1991) p178 McGraw Hill. New York.

[60]Petty RE, Cacioppo JT. Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *Journal of Personality and Social Psychology.* 1979;**37**:1915-1926.

[61]Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. In Bekowicz L (Ed.) *Advances in experimental psychology.* 1986;**19**:123-205.

[62]Chaiken S. Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology.* 1980;**39(5)**:752-766.

---

[63]Petty RE. Cacioppo JT. The elaboration likelihood model of persuasion. In Berkowitz L Ed. *Advances in experimental social psychology.* 1986:**19**:123-205. Academic Press. New York.

[64]Zimbardo PG, Leippe MR. The psychology of attitude change and social influence (1991) McGraw Hill. New York.

[65]Barone DF, Hutchings PS. Cognitive elaboration: Basic research and clinical application. *Clinical Psychology Review* . 1993; **13**:187-201.

[66].Hume D. (Selby-Bigge LA). A Treatise of Human Nature (1976) p89. Oxford University Press. Oxford.

[67] Skinner BF. Science and human behaviour. 1953. Macmillan.

[68] Brook RH, Williams KN. Effect of medical care review on the use of injections. A study of the New Mexico experimental medical care review organisation. *Ann. Intern. Med.* 1976;**85**:509-15.

[69]Baker R. Audit and standards in new general practice. *BMJ* 1991;**303**:32-4.

[70] A prescription for improvement. Audit Commission. 1994. HMSO.

[71] Statistical Bulletin 1993/8. Statistics of prescriptions dispensed in the Family Health Services Authorities: England 1982 to 1992. Department of Health.

[72] Bateman DN, Campbell M, Donaldson LJ, Roberts SJ, Smith JM. A prescribing incentive scheme for non-fundholding general practices: an observational study. BMJ, 1996;313:535-8.

---

[73] Kelsey MC, Kouloumas GA, Lamport PA, Davis CL. Relation between general practitioners' prescribing of antibacterial drugs and their use of laboratory tests. BMJ, 1996;313:922.

[74] Winkens RAG, Pop P, Grol RPTM, Kester ADM, Knottnerus JA. Effect of feedback on test ordering behaviour of general practitioners. BMJ, 1992;304:1093-6.

[75] Winkens RAG, Pop P, Bugter-Maessen AMA, Grol RPTM, Kester ADM, Beusmans GHMI, Knottnerus JA. Randomised controlled trial of routine individual feedback to improve rationality and reduce numbers of test requests. Lancet, 1995;345:498-501.

[76]Buntinx F, Knottnerus JA, Crebolder HFJM, Seegers T, Essed GGM, Schouten H. Does feedback improve the quality of cervical smears? A randomised controlled trial. *British Journal of General Practice*. 1993;**43**:194-198.

[77]Winkens RAG, Pop P, Grol RPTM, Kester ADM, Knottnerus JA. Effect of feedback on test ordering behaviour of general practitioners. *BMJ.* 1992;**304**:1093-6.

[78]Avorn J, Soumerai SB. Improving drug-therapy decisions through educational outreach. *New England Journal of Medicine*.1983;**308**:1457-63.

[79]Soumerai SB, Avorn J. Economic and policy analysis of university-based drug "detailing". *Medical Care*. 1986;**24**:313-331.

---

[80]Newton-Syms FAO et al. The influence of an academic representative on prescribing by general practitioners. *British Journal of Clinical Pharmacology.* 1992;**33**:69-73.

[81] Kelly FJ, Kelly HM. What they really teach you at the Harvard Business School. 1987. Grafton.

[82] Newble D, Cannon R. A handbook for medical teachers.1994. Kluwer.

[83] Handy C. Understanding organisations. 1990. Penguin.

[84]Green PE. The general practice formulary - its role in rational therapeutics. *Journal of the Royal College of General Practitioners.* 1985;**35**:570-2.

[85]Beardon PHG, Brown SV, Mowat DAE et al. Introducing a drug formulary to general practice - effects on practice prescribing costs. *Journal of the Royal College of General Practitioners.* 1987;**37**:305-7. In Gilleghan JD. Prescribing in General Practice. (1991) p18 The Royal College of General Practitioners. London.

[86]Van Zwanenberg TD, Grant GB, Gregory DA. Can rational prescribing be assessed? *Journal of the Royal College of General Practitioners.* 1987;**37**:308-10.

[87]McGavock H. Promoting rational therapeutics: the Swedish guidelines and a UK response. *British Journal of General Practice.* 1990;**40**:47-8.

[88]McCarthy M, Wilson-Davis K, McGavock H. Relationship between the number of partners in a general practice and the number of

different drugs prescribed by that practice. *British Journal of General Practice.* 1992;**42**:10-12.

[89]Branthwaite A, Ross A, Henshaw A, Davie C. Continuing education for general practitioners. Occasional paper 38. 1988, p17. The Royal College of General Practitioners. London

[90]Xamoterol (Corwin) - Revised indications, contraindications, dose schedule and warnings. (1990) *Current Problems* 28. Committee on Safety of Medicines. London.

[91]Plumridge RJ, Berbatis CG. Drug bulletins: effectiveness in modifying prescribing and methods of improving impact. *The Annals of Pharmacotherapy, DICP.* 1989;**23**:330-34.

[92] Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. In Bekowicz L (Ed.) *Advances in experimental psychology.* 1986;**19**:123-205

[93]Smith R. Doctors and the drug industry: too close for comfort. *BMJ.* 1986;**293**:905.

[94]Gilleghan JD. Prescribing in general practice. Occasional paper 54. 1991, p15. The Royal College of General Practititoners. London.

[95]Peay MY and Peay ER. The role of commercial sources in the adoption of a new drug. *Social Science and Medicine.* 1988;**26**:1183-89.

[96]Taylor RJ. Bond CM. Change in the established prescribing habits of general practitioners: an analysis of initial prescriptions in general practice. *British Journal of General Practice.* 1991;**41**:244-48.

---

[97]Kiesler CA. The psychology of commitment: Experiments linking behaviour to belief. 1971. Academic Press. New York.

[98]Chapman S. Sales and marketing skills for professional advisers. (1992) West Midlands Regional Health Authority.

[99]Zimbardo PG, Leippe MR. the psychology of attitude change and social influence. (1991) p80. McGraw Hill. New York.

[100]Jack DB, Soppitt AJ. Give a drug a bad name. *BMJ.* 1991;**303**:1606-8.

[101]Miller GA. the magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review.* 1956;**63**:81-97.

[102]Dieppe PA, Frankel SJ, Toth B. Is research into the treatment of osteoarthritis with non-steroidal anti-inflammatory drugs misdirected? *Lancet.* 1993;**341**: 353-4.

[103]Horton R. Parmacoeconomical with the truth. *Lancet.* 1993;**341**:752.

[104]Waller PC, Wood SM, Langman MJS, Breckenridge AM, Rawlins MD. Review of company postmarketing surveillance studies. *BMJ.* 1992;**304**:1470-2.

[105]Hayes TM, Allery LA, Harding KG, Owen PA. Continuing education for general practice and the role of the pharmaceutical industry. *British Journal of General Practice.* 1990;**40**:510-12.

[106]Sheldon TA, Davey Smith G. Consensus conferences as drug promotion. *Lancet.* 1993;**341**:100-102.

---

[107] Taylor RJ, Bond CM. Change in the established prescribing habits of general practitioners: an analysis of initial prescriptions in general practice. BJGP, 1991;41:244-8.

[108] Fisher A. The logic of real arguments. (1988) Cambridge University Press. Cambridge.

[109] Eddy D. Clinical Decision-making: From theory to practice. Practice policies - what are they? *Journal of the American Medical Association* 1990;**263**:877-80.

[110] Goodwin P, Wright G. Decision analysis for management judgement. 1991. Wiley. Chichester.

[111] Makridakis S, Wheelwright SC. Forecasting methods for management. 5th Edition. 1989. Wiley.Chichester.

[112] Onion CWR. Patient empowerment. *Medicines, Economics and Health*. In press.

[113] Bolam v Friern Hospital Management Committee (1957) 2 AER 118.

[114] Anonymous. *General Practitioner*. March 12 1993 p44-5.

[115] Hunter v Hanley (1955) SC 200

[116] Hart. Protocols or guidelines or neither? *British Medical Journal*. 1993;**306**:816.

[117] Marshall JH. *Oxford English Dictionary Word and Language Service*. Personal correspondence. 5th May 1993.

[118] Hillman AL, Pauly MV, Kerstein JJ. How do financial incentives affect physicians' clinical decisions and the financial performance of

health maintainance organisations? *N Engl J Med* 1989;**321**:86-92.

[119]Stringer MD, Capps SNJ. Rationalising the management of swallowed coins in children. *BMJ* 1991;**302**:1321-2.

[120]Fotherby MD, Harper GD, Potter JF. General practitioners' management of hypertension in elderly patients. *BMJ* 1992;**305**:750-2.

[121]Davey Smith G, Pekkanen J. Should there be a moratorium on the use of cholesterol lowering drugs. *BMJ* 1992;**304**:431-4.

[122]Jones R, Rawlins MD. Prescribing at the interface between hospitals and general practitioners. *BMJ* 1992;**304**:4-5.

[123]Chisholm R. Guidelines for radiological investigations. *BMJ* 1991;**303**:797-8.

[124]Farmer A. Medical practice guidelines: lessons from the United States. *BMJ* 1993;**307**:313-7.

[125]Farmer A. Setting up consensus standards for the care of patients in general practice. *British Journal of General Practice*. 1991:135-6.

[126]Jenkins D. Investigations: how to get from guidelines to protocols. *BMJ*. 1991;**303**:323-4.

[127]Jackson MN, Nutting PA. Clinical guidelines development: Opportunities for family physicians. *The Journal of Family Practice*. 1991;**33:2**:129-32

---

[128]Onion CWR, Bartzokas CA. Microbes: reference data for the management of infections in the Wirral community. 1993 Wirral Hospital NHS Trust.

[129]Eddy DM. A Manual for Assessing Health Practices and Designing Practice Policies: The Explicit Approach. 1992 The American College of Physicians. Philadelphia.

[130]Linsk JA. Clinical decision making: Theory vs practice. Letter *JAMA* 1990;**264**:12:1533.

[131]Paykel ES, Priest RG. Recognition and management of depression in general practice: Consensus statement. BMJ 1992;305:1198-1202.

[132]Joint Working Group of the Research Unit of the Royal College of Physicians and the British Paediatric Association. Guidelines for the management of convulsions with fever. *BMJ* 1991;**303**:634-6.

[133]Sever P et al. Management guidelines in essential hypertension: report of the second working party of the British Hypertension Society. *BMJ.* 1993;**306**:983-7.

[134]Bayliss et al. GP management guidelines in CHF. 1992 Bristol-Myers Squibb Pharmaceuticals Ltd.

[135]Chisholm G et al. Benign Prostatic Hyperplasia: the Shared-Care Initiative. 1993 RRAssoc, Chineham, Hants. Sponsored by Abbott Laboratories.

[136]*Effective Health Care.* Reports 1993 Onwards. The Department of Health. Leeds.

[137]Russell I, Grimshaw J, in Roland M, Coulter Eds. Hospital Referrals. 1992 Oxford University Press. Oxford. p201.

[138]Skrabanek P. The epidemiology of errors. *Lancet* 1993;**342**:1502.

[139]Eysenck HJ. Meta-analysis and its problems. BMJ. 1994;309:789-92.

[140]Feinmann J. Can GPs trust the DoH economists? *General Practitioner* 19th November 1993.

[141]Carr-Hill RA, Morris J. Current practice in obtaining the "Q" in QALYs: a cautionary note. *BMJ* 1991;**303**:699-701.

[142]Harris J. More and better justice. In Bell JM, Mendus S Eds. Philosophy and Medical Welfare. Royal Institute of Philosophy lecture series; 23. 1988. Press Syndicate of the University of Cambridge.

[143]Royal College of Radiologists Working Party. Influence of Royal College of Radiologists' guidelines on referral from general practice. *BMJ* 1993;**306**:110-1.

[144]Buntinx et al. Does feedback improve the quality of cervical smears? A randomised controlled trial. *British Journal of General Practice* 1993;**43**:194-198.

[145]Zimbardo PG, Leippe MR. The Psychology of Attitude Change and Social Influence. 1991 McGraw Hill. New York. p169.

[146]Russell IT et al. North of England Study of Standards and Performance in General Practice. Medical audit in general practice. I: Effect on doctors' clinical behaviour for common childhood conditions. *BMJ* 1992;**304**:1480-4.

[147]Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet* 1993;**342**:1317-22.

[148]Onion CWR, Chm. Minutes of the first meeting of the Wirral Medical Management Guidelines editorial board. 16th January 1992. Wirral Family Health Services Authority. Birkenhead.

[149]Eddy D. Practice policies: where do they come from? *JAMA* 1990; **263**:1265-75.

[150]Zimbardo PG, Leippe MR. The psychology of attitude change and social influence. 1991. McGraw Hill. New York.

[151]Osgood CE, Suci GJ, Tannenbaum PG. The measurement of meaning. 1957. Urbana: The University of Illonois Press.

[152]Grol R. National standard setting for quality of care in general practice: attitudes of general practitioners and response to a set of standards. *British Journal of General Practice*, 1990, **40**, 361-364.

[153]Fazio RH, Zanna MP. Direct experience and attitude-behaviour consistency. In Berkowitz L Ed. *Advances in experimental social psychology*. 1981;**14**:162-202. New York Academic Press.

---

[154]Dunnell K, Cartwright A. Medicine Takers, Hoarders and Prescribers. 1972. Routledge and Kegan Paul. London.

[155]Fry J. Common Diseases: Their nature incidence and care. 1979. MTP. London.

[156] Kuhn TS. The structure of scientific revolutions. 1962. The University of Chicago Press.

[157]Popper KR. Objective knowledge: An evolutionary approach. 1979. Oxford University Press.

[158]Copi IM. Symbolic Logic.1979. MacMillan.

[159]Peckham M. Research for Health. 1993. Department of Health.

[160]Branthwaite A, Ross A, Henshaw A, Davie C. Continuing education for general practitioners. Occaisional paper 38. 1988. The Royal College of General Practitioners. London.

[161]Avorn J, Soumerai SB. Improving drug-therapy decisions through educational outreach. *New England Journal of Medicine.* 1983;**308**:1457-63.

[162]England AG. Information to doctors about medicines:supporting the advisers. In: Wells FO Ed. Medicines: Responsible Prescribing. 1992. The Queens University of Belfast.

[163]Roethlisberger FJ, Dickson WV, Management and the worker: an account of a research program conducted by the Western Electric Co., Hawthorne Works, Chicago. 1939. Harvard University Press.

---

[164]Saunders C, Stringer M. Rewarding People: The Skill of Responding Positively. 1993. Routledge. London.

[165]Brazier M. Medicine, Patients and the Law. 1987. Penguin. p73.

[166]Hawkins C. Mishap or Malpractice? 1985. Medical Defence Union/Blackwell Scientific. p169.

[167]Onion CWR, Walley T. Clinical guidelines: development, implementation , and effectiveness. *Postgrad Med J* 1995;**71**:3-9.

[168]Onion CWR, Bartzokas CA. Microbes: reference data for the management of infections in the wirral community. 1993. Wirral Health and Wirral Hospital (NHS) Trust.

[169]Patient Care Information System (PCIS). Wirral Medical Microbiology. Wirral Hospital (NHS) Trust. 1992-3.

[170]Wilkie ME, Almond MK, Marsh FP. Diagnosis and management of urinary tract infection in adults. British Medical Journal. 1992;**305**:1137-41.

[171]Lewis DA, Reeves DS. Antibacerial agents: trimethoprim and co-trimoxazole. *Prescribers' Journal.* 1995;**35(1)**:25-31.

[172] Revised indications for co-trimoxazole (Septrin, Bactrim, etc.). *Current Problems in Pharmacovigilance.* Committee on Safety of Medicines, Medicines Control Agency. 1995;**21**:6.

[173]Fry J. Common Diseases: Their Nature, Incidence and Care. 1979. MTP Press.

[174]In-vitro activity of anti-bacterial drugs and clinical practice. *Drug and Therapeutics Bulletin.* 1995;**33(9)**:71-2.

[175]Bradley CP. Uncomfortable prescribing decisions: a critical incident study. *BMJ.* 1992. **304**:294-6.

[176]Fleming DM, Norbury CA, Crombie DL. Annual and Seasonal Variation in the Incidence of Common Diseases. Occaisional Paper 53. 1991. The Royal College of General Practitioners. London.

[177]Davey P, Rutherford D, Graham B, Lynch B, Malek M. Repeat consultations after antibiotic prescribing for respiratory infection: a study in one general practice. *British Journal of General Practice.* 1994;**44**:509-513.

[178]Campbell MJ, Machin D. Medical Statistics: A Commonsense Approach. 1990. Wiley.

[179]Forster DP, Frost CEB. Use of regression analysis to explain the variation in prescribing rates and costs between family committees. *British Journal of General Practice.* 1991;**41**:67-71.

[180]Gore SM. Altman DG. Statistics in Practice. 1982. British Medical Association.

[181]Swinscow TDV. Statistics at Square One. 1989. British Medical Association.

[182]Porkess R. Dictionary of Statistics. 1988. Collins.

---

[183]The Third National Morbidity Survey. OPCS, DoH, RCGP. 1981. HMSO.

[184]Pocock SJ. Clinical Trials: A Practical Approach.1983. Wiley.

[185]Onion CWR, Slade PD. Depth of information processing and memory for medical facts. *Medical Teacher.* 1995;**17(3)**:307-313.

[186]Onion CWR, Dutton CE, Walley T, Dunne WH, Turnbull CJ, Buchan IE. Clinical guidelines: description and evaluation of a method for their local development and implementation. *Family Practice.* 1996;13(1):28-34.

[187] Busch W. Schein und sein. 1909, reproduced in Popper KR. Objective knowledge. 1972. Clarendon Press. Oxford.

[188]Gore SM, Altman DG. Statistics in practice. 1982. BMA

[189] Honey P, Mumford A. The manual of learning styles. 1992. Honey. Maidenhead.

[190] Guglielmino LM. Development of the Self-Directed Learning Readiness Scale. (Doctoral dissertation, University of Georgia) Dissertation Abstracts International. 1977;38:6467A.

[191] Bligh J. The S-SDLRS: a short questionnaire about self-directed learning. Postgraduate Education for General Practice. 1993;4:121-125.

[192]Onion CWR, Bartzokas CA. Microbes: reference data for the management of infection in the Wirral community. 1993. Wirral Hospital (NHS) trust.

---

[193]Newble D, Cannon R. A handbook for medical teachers. 3rd edition. 1994. Kluwer. London.

[194]Petty RE, Caccioppo JT. The elaboration likelihood model of persuasion. In Berkowitz L. Ed. *Advances in experimental social psychology*. 1986;**19**:123-205. Academic Press. New York.

[195]Ford ME. Motivating humans: Goals, emotions and personal agency beliefs. 1972.Sage. London.

[196]Bradley CP. Uncomfotable prescribing decisions: a critical incident study. *British Medical Journal.* 1992;**304**:294-6.

[197]Audit Commission. A prescription for improvement. 1994. HMSO.

[198]Zimbardo PG, Leippe MR. The psycholgy of attitude change and social influence. 1991. McGraw Hill. New York.

[199] Quine WVO. From a logical point of view. 1952. Harvard University Press.

[200]Cartwright A.Health surveys in practice and in potential. 1986. King's Fund.

[201] Gart JJ, Nam J. Approximate interval estimation of the ratio of bionomial parameters: a review and corrections for skewness. *Biometrics.*1988;**44**:323-338.

[202]Armitage P, Armitage G, Berry G. Statistical methods in medical research. 1987. Blackwell.

[203]Gardner MJ, Altman DG. Statistics with confidence - confidence intervals and statistical guidelines. 1989. BMJ.

---

[204]Statistical bulletin 1995/15: statistics of prescriptions dispensed in the Family Health Services Authorities: England 1993 to 1994. July 1995. HMSO

[205]Eypasch E, Lefering R, Kum CK, Troidl H. Probability of adverse events that have not occurred yet: a statistical reminder. *BMJ* 1995;**311**:619-20.

[206]Harris CM, Heywood PL, Clayden AD. The analysis of prescribing in general practice: a guide to audit and research.1990. HMSO.

[207]Prescription Pricing Authority Annual Report April 1989 - March 1990. HMSO.

[208]Morbidity statistics from general practice. Fourth national study 1991-1992. McCormick A, Fleming D, Charlton J. 1995 HMSO.

[209] Kuhn TS. The structure of scientific revolutions. 1962. The University of Chicago Press.

[210]Tompkins DS, Shannon A-M. Clinical value of microbiological investigations in general practice. *British Journal of General Practice.* 1993;**43**:155-158.

[211]Humphris D. Clinical guidelines: an industry for growth. *Nursing Times.* 1994;**90:40**:46-47.

[212]Conroy M, Shannon W. Clinical guidelines: their implementation in general practice. *British Journal of General Practice.* 1995;**45**:371-375.

[213]Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet.* 1993;**342**:1317-1322.

[214]Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. In Berkowitz L (Ed) *Advances in Experimental Psychology* 1986;**19**:123-205. Academic Press. New York.

[215]Roland MO, Middleton J, Goss B, Moore AT. Should performance indicators in general practice relate to whole practices or to individual doctors? *Journal of the Royal College of General Practitioners.* 1989;**39**:461-462.

[216]Bogle SM, Harris CM. Measuring prescribing: the shortcomings of the item. *British Medical Journal.* 1994;**308**:637-40.

[217]Statistical bulletin: statistics of prescriptions dispensed in the Family Health Services Authorities: England 1984 to 1994. Government Statistical Service. 1995. HMSO.

[218]A prescription for improvement: towards more rational prescribing in general practice. Audit Commission. 1994. HMSO

[219]Harris CM, Heywood PL, Clayden AD. The analysis of prescribing in general practice: a guide to audit and research. 1990. HMSO.

[220]The development and implementation of clinical guidelines. Report from General Practice 26. 1995. The Royal College of General Practitioners.

---

[221]Effective health Care. Implementing clinical practice guidelines. 1994;**8.** University of Leeds.

[222]Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet* 1993;**342**:1317-22.

[223]Yoong AFE, Lim J, Hudson CN, Chard T. Audit of compliance with antenatal protocols. *BMJ* 1992;**305**:1184-6.

[224]Royal College of Radiologists Working Party. A multicentre audit of hospital referral for radiological investigation in England and Wales. *BMJ* 1992;**304**:740-3.

[225]Maiman LA, Greenland P, Hildreth NG, Cox C. Patterns of physicians' treatments for referral patients from public cholesterol screening. *Am J Prev Med* 1991;**7(5):**273-9.

[226]Lane DS, Polednak AP, Burg MA. Effect of continuing medical education and cost reduction on physician compliance with mammography screening guidelines. *The Journal of Family Practice* 1991;**33(4):**359-368.

[227]British Thoracic Society and others. Guidelines on the management of asthma. *Thorax* 1993;**48**(supplement):S1-24.

[228]Seabrook M, Booton P, Evans T. Widening the horizons of medical education. 1994 Kings Fund Centre.

[229] Kuhn TS. The structure of scientific revolutions. 1962. The University of Chicago Press.

---

[230] Onion CWR, Slade PD. Deep processing and memory for medical facts. *Medical Teacher.* 1995;**17(3)**:307-313.

[231] Onion CWR, Dutton CE, Walley T, Turnbull CJ, Dunne W, Buchan IE. Local clinical guidelines: description and evaluation of a participative method for their development and implementation. *Family Practice.* 1996;13(1):28-34.

[232] Kuhn TS. The structure of scientific revolutions. 2nd edition. 1970. University of Chicago Press.

[233] Anon. Protocol helps GPs decide best use of ACE inhibitors. General Practitioner. 22nd April 1994.

[234] Anon. Clinical guidelines cause concern to anxious GPs. *BMA News Review.* 14th June 1995.

[235] Feder G. Clinical guidelines in 1994: let's be careful out there. *British Medical Journal.* 1994;**309**:1457-8.

[236] Honigsbaum F. Rationing of healthcare. *NAHAT update.* August 1994;**6**:5.

[237] Charlton BG. Practice guidelines and practical judgement: the role of mega-trials, meta-analysis and consensus. *British Journal of General Practice.* 1994;**44**:290-291.

[238] Baker R, Fraser RC. Development of review criteria: linking guidelines and assessment of quality. *British Medical Journal.* 1995;**311**:370-371.

---

[239]Dans PE. Credibility, cookbook medicine, and common sense: guidelines and the College. *Annals of Internal Medicine.* 1994;**120:11**:966-967.

[240]Bligh J. Problem based, small group learning. *British medical Journal.* 1995;**311**:342-343.

[241]Bazzoli F, Zagari RM, Fossi S *et al.* Efficacy and tolerability of a short-term, low dose triple therapy for eradication of *Helicobacter pylori.* Gastroenterology. 1993;**104**:4:A40.

[242]Axon A. Drug regimens. In: Northfield TC, Mendall M, Goggin PM (Eds.) Helicobacter pylori infection. 1993. Kluwer.

[243]Mill JS. On Liberty. 1974. Penguin.

[244]Bligh J. MD thesis. 1994 University of Liverpool.

[245]Ford ME, Nichols CW. Manual: Assessment of personal goals. 1992. Palo Alto CA: Consulting Psychological Press.

[246]Seabrook M, Booton P, Evans T. Widening the horizons of medical education. 1994 Kings Fund Centre.

# Appendices

# Appendix A

# Critical appraisal of some relevant published studies

Fourteen recent studies, selected because they address some of the issues explored in the thesis, are reviewed to display the benefits and problems associated with methodologies used by other researchers active in the field. Critical appraisal of other authors' work has been separated from the main text in order to allow uninterrupted logical progression in the thesis.

## Contents

## Variations in practice

### Baker R. General practice in Gloucestershire, Avon and Somerset: explaining variations in standards. BJGP 1992;42:415-8

*Purpose of study*

To explore the reasons for variation in the level of development in local general practices.

*Population*

All 324 practices in the above Family Health Services Authorities (FHSAs).

*Method*

Questionnaire survey, comprehensive - all aspects of quality in general practice, 72 main questions, piloted in 1982.

*Data*

Response rate = 89%. Two exclusions. No power calculation. 'Development scores' calculated for each practice.

*Analysis*

Multiple regression analysis. Median score was 52 out of 95, range 14 to 74. The important factors associated with 'high quality' practices were:

- Training practice (trainee)
- Practice manager
- Younger practitioners
- Large list sizes
- Low underprivileged (Jarman) score

These factors statistically explained 42% of the variation.

*Validity of conclusion*

The authors draw the conclusion that addressing the above factors (e.g. taking on a practice manager or becoming a training practice) will cause improved quality. However, it is just as likely that well motivated, up to date GPs would seek to become trainers and to delegate appropriate duties to appropriately recruited administrative staff. The authors have identified useful indicators of quality practices, but have not proved causation.

*Overall appraisal*

An ambitious and comprehensive project well executed. The quality indicators were not weighted so some factors may have had an effect on scores disproportionate to their true importance.

Quality is highly subjective, some evidence of validation, by interested parties (other doctors, patients, FHSA managers etc.), of the parameters measured would have been reassuring.

*Relevance to the thesis and main study*

The study supports regarding training practice status as an important confounding variable (chapter 6 in the thesis) and raises concerns about the wide variation in quality amongst practices (chapter 1).

## Forster DP, Frost CEB. Use of regression analysis to explain the variation in prescribing rates and costs between family practitioner committees. BJGP, 1991;41:67-71

*Purpose of study*

To explore one method of predicting or explaining observed variations in prescribing rates and costs between Family Practitioners Committees (FPCs - fore-runners to FHSAs) in terms of need for health care, including age-sex structure and indicators of resource availability.

*Population*

All 98 FPCs in England and Wales.

*Method*

Survey of nationally held data for each FPC.

*Data*

Prescribing data (total cost and total number of prescriptions dispensed) drawn from 1 in every 200 prescriptions from (dispensing and non-dispensing) GPs, standardised mortality ratios, demographic structure, underprivileged scores and number of GPs in each FPC in 1987.

*Analysis*

Multiple regression analysis.

Median number of prescriptions per patient in 1987 for FPCs was 7.7 (range 5.7 to 11.1). Median cost per patient was £41.9p (range £35.00p to £ 54.1p). Sixty percent of the variation in costs and 65% in rate were explained by SMR (all causes) and GP supply (GPs per capita). Underprivileged scores were not significantly associated with variance.

Wirral FPC was not in the extreme centiles of prescribing rate or cost.

*Validity of conclusion*

Authors conclude that prescribing rate and cost is proportional to need (as expressed by SMR all causes) and resource availability (as expressed by GP supply per capita), however, >40% of variation remained unexplained. Their assertion that resource allocation for prescribing based upon this data would be likely to omit the effect important unknown factors so leading to inequity seems to be justified.

*Overall appraisal*

A meticulous study of the potential for a model for prescribing resource allocation based upon currently, nationally held data on indicators of need and resource. The analytic method was appropriate and the conclusions valid.

*Relevance to the thesis and main study*

Describes wide prescribing variation despite aggregation of the data into FPCs (around 100 practices per FPC) - as discussed in chapter 1. Demonstrates the importance of local mortality as an indicator of a morbidity confounding variable, supporting the

balance of geographical localities (differing SMRs) during the selection process for the main study (chapter 65). The fact that Wirral FPC did not figure as a prescribing outlier validates the universality of the main study (chapter 9).

## Prescribing

### Bradley CP. Uncomfortable prescribing decisions: a critical incident study. BMJ, 1992;304:294-6

*Purpose of study*

To determine (a) whether discomfort in prescribing is related to particular drugs or to particular patient characteristics (e.g. morbidity) and (b) the reasons for decisions and why they felt uncomfortable.

*Population*

A sample of 69 (51%) from all 168 GPs in a north England FPC and 5 trainees. They came from a "wide range of ages, backgrounds, and practice situations, though younger doctors, women doctors, doctors with higher qualifications, and doctors in practices with more partners and smaller list sizes were slightly over-represented". No description of method of sample selection. Three doctors interviewed were excluded - no reasons.

*Method*

Semi-structured interview. GPs asked to record or recall incidents where their prescribing made them feel 'uncomfortable' over a period of 10 days prior to interview.

*Data*

Categorisation of mentioned classes of drugs, disease groups, reasons for decisions made, and reasons for discomfort.

*Analysis*

Three hundred and seven incidents recorded by the 74 doctors (mean 4.1). Simple rankings of proportions in each type of data. Commonest (modal) class of drug was antibiotics (23%), commonest disease group was respiratory tract infection (23%), commonest reason for decision was patient expectation (45%) and commonest reason for discomfort was concern about drug toxicity (27%).

*Validity of conclusion*

The author concludes that many considerations, including medical, social and logistical (as well as pharmacological appropriateness, safety, effectiveness and economy) influence the decision to prescribe in general practice. This does not reflect the

purpose of the study, and is an extrapolation from this study of specifically unusual prescribing incidents to general prescribing decisions.

*Overall appraisal*

It is possible that the sample is biased, there is no description of the sample selection method, nor of exclusion criteria. The structured interview approach is highly appropriate to explore influences on prescribing decisions. Specifically targeting areas where decisions have gone wrong is a good way to learn where education needs to be directed to prevent such experiences. However, such a study can only imply possible, not probable, influences in ordinary prescribing. The notion of 'feeling uncomfortable' is vaguely defined so different GPs may have been discussing incidents incurring different sensations A similar paper by the same author later in the year (1992) paper examining the same data (Bradley CP. Factors which influence the decision whether or not to prescribe: the dilemma facing general practitioners. BJGP, 1992;42:454-8) adds a little more analysis, and expressed 'discomfort' in terms of cognitive dissonance - a term proposed by Festinger in 1957 to refer to internal inconsistency between different cognitions in the mind of a person, for example when a person decides to act in a way he

believes to be unsatisfactory - i.e. attitude / behaviour inconsistency.

The data is clearly presented but no examination of the number or type of factors presented by different sub-groups of the sample is displayed; there is no statistical testing of associations, rank correlation or analysis of proportions - perhaps because of insufficient data. The study does however, expose the complexity of factors that influence 'uncomfortable' prescribing decisions and offers a unique insight into the workings of prescribers' minds.

*Relevance to the thesis and main study*

The finding that antibiotics were the modal class of drugs, and respiratory tract infection the modal disease group, causing concern confirms infections as a suitable topic area for the main study (chapter 6). If doctors are assumed to feel uncomfortable where their goals in a consultation have been frustrated then the nature of doctors' goals can be deduced (chapter 1), this supports the 'alignment of goals' component of the agenda design for the deep processing visits (chapter 6).

**Taylor RJ, Bond CM. Change in the established prescribing habits of general practitioners: an analysis of initial prescriptions in general practice. BJGP, 1991;41:244-8**

*Purpose of study*

To describe (a) the types of drugs prescribed by GPs in a large sample of initial prescriptions, (b) the additions and deletions made to the doctors' repertory, and (c) the factors influencing the change.

*Population*

All 281 GPs in the Grampian region.

*Method*

Prospective survey of 100 consecutive initial prescriptions (intended as a complete course of treatment or the first in a series of prescriptions for the same drug) every 7½ weeks for 1 year (1985 / 6), each labelled accordingly by the prescriber as his or her habitual choice, a drug he or she adopted in the last year, or a drug dropped from his or her personal repertory in the last year but prescribed on this occasion for some reason. Where there was a change in repertory the prescriber was asked to identify the

influence(s) for that change. All the data was collected on special prescription forms. A time series technique (staggered recording) in prescription sampling protected against confounding by seasonal variation.

*Data*

Two hundred and twelve (75%) of GPs took part. There was a slight over-representation of members of the Royal College of General Practitioners and 23 GPs were excluded.

*Analysis*

The data was categorised by therapeutic (British National Formulary) grouping, newly adopted drugs, and sources of influence. The results were expressed as ranked proportions and pie diagrams; no further statistical analysis was presented.

The proportion of drugs added to repertory in the previous 12 months was 5.4% of all initial prescriptions. Less than 1% had been drugs dropped from the usual repertory in the last 12 months. The commonest therapeutic group of newly adopted drug was central nervous system (11%) followed closely by infections (16.3%). Three antibiotics, augmentin (replacing amoxil), co-trimoxazole (replacing septrin) and erythrocin (replacing erythroped) were in the top 6 most frequently prescribed newly

adopted drugs. Commonest influences for changes in the repertory were said to be 'limited list' regulations (25%), pharmaceutical company representative (20%) and hospital doctor (18%) - postgraduate education was not mentioned. The commonest initial drug group was antibiotics (34%). The relative proportions of different antibiotics (generally complete courses in each prescription and therefore nearly always 'initial') were compared with all Scotland figures to detect any Hawthorne effect; proportions were similar.

*Validity of conclusion*

The authors conclude that (i) GPs were not unduly affected by commercial sources of information and (ii) their prescribing habits were stable and conservative. If the limited list (a current unique factor) is ignored, the greatest influence on adoption of a drug was the pharmaceutical company representative, so (i) is clearly an illogical conclusion. For all courses of drugs initiated in the study period, more than 5% were newly adopted in the previous 12 months; whether this represents conservatism and a stable situation is a matter for the reader's subjective opinion, but assertion (ii) is certainly debatable.

*Overall appraisal*

The response rate was good, the data collection method ingenious, prevention of obscuring by repeat prescribing complete, data validation rigorous and the definitions clear. Analysis was simple and effective, but statistical associations between the different categories of data were not explored (e.g. which type of newly adopted drug was influenced most by commercial information?). The aims of the study were achieved, but the conclusions drawn seem somewhat inconsistent with the findings.

*Relevance to the thesis and main study*

The evidence supports concerns that postgraduate medical education is not perceived to be effective (chapter 1). Contrary to the authors' conclusion, the drug rep. is considered the most important influence on the adoption of new drugs by GPs, this supports the argument in chapter 2. The most commonly prescribed type of initial drug is the antibiotic, suggesting that infections are a most suitable area for testing an initiative to change prescribing, as with the main study (chapter 6).

**McCarthy M, Wison-Davis K, McGavock H. Relationship between the number of partners in a general practice and the number of different drugs prescribed by that practice. BJGP, 1992;42:10-12**

*Purpose of study*

To ascertain whether practice size, as measured by number of doctors, has any bearing on the range of drugs prescribed.

*Population*

All 132 practices in Northern Ireland.

*Method*

Survey of prescribing in all practices in the month of January 1989. Synchronous short questionnaire survey to each practice to ascertain whether they used a drug formulary or had a verbally agreed prescribing policy.

*Data*

All NHS prescriptions in the province dispensed in the study month were collected from the national database. The range of prescribing was assessed by counting the number of preparations in each of 22 different therapeutic groups prescribed by each

practice. Each therapeutic group was sub-divided in symptomatic, systematic or intermediate classes. Response rate for the questionnaire was 72%.

*Analysis*

Twenty-two percent of practices were single-handed, the rest ranged from 2 to 7 partners (median {not stated} =3 partners). Mean list size per doctor was 1844 (range 1033 - 3363).

There was a strong positive correlation (r=0.85) between the number of doctors in a practice and the range of drugs prescribed. There was no correlation (r=0.16) between drug cost per patient and range of drugs prescribed.

Multiple regression analysis indicated that the number of practitioners in a practice statistically explained 71% of the variation in range, list size could account for only 5% of the variation. There was no difference between symptomatic, systematic and intermediate drugs classes.

Possession of a formulary was declared by 20% of respondents, 34% had a verbally agreed policy and 46% had no policy. The use of a formulary or agreed policy had no significant association with the range of drugs prescribed by a practice, nor any associated reduction in prescribing costs.

*Validity of conclusion*

The authors conclude that the use of a practice prescribing policy had no effect on prescribing range or cost and infer that that therapeutic policies may be difficult to implement within group practices. However, the inference is vacuous. A single-handed practitioner must have 100% agreement on prescribing and it is not surprising that disagreement is found with multiple partners. The fact that a practitioner prescribes a narrow range of drugs does not mean that he is adhering to a planned therapeutic policy.

*Overall appraisal*

An important paper that exploded several myths - it transpires that formularies do not reduce the range or cost of prescribing; peer pressure from partners does not lead to a narrow range of prescribed drugs and; a narrow range is not associated with reduced prescribing cost per patient. However, range of drugs prescribed is a poor indicator of good prescribing or compliance with a formulary. Direct comparison of individual drugs prescribed with those in a formulary carefully developed for general practice would be a better measure. The study suffers from the usual problem of studies utilising anonymised prescription data, it is possible that apparently well prescribed 'good' drugs are being prescribed for inappropriate conditions -

this study compounds the problem by using range, an even more remote indicator.

*Relevance to the thesis and main study*

Formularies, by themselves, do not influence prescribing (chapter 2). Single-handed practices are different to other GPs in their prescribing behaviour and so should be balanced across the groups in the main study (chapter 6).

## Fundholding and prescribing incentive schemes

### Harris CM, Scrivener G. Fundholders' prescribing costs: the first five years. BMJ, 1996;313:1531-4

*Purpose of study*

To compare successive waves of fundholding general practices with continuing non-fundholders with regard to prescribing rate and cost.

*Population*

All practices in England in the 6 years from April1990 to March1996.

*Method*

Cohort study of changes and rates of change in cost (net ingredient cost) and changes in number of items per age-weighted head of population in fundholding practices, before and after fundholding, relative to non-fundholders.

*Data*

Quarterly cost and rate information for each practice from a database of all NHS prescriptions dispensed in England held by

the Prescription Pricing Authority. The cumulative total of fundholders was 2649. The <10 practices that ceased fundholding in the study period were not excluded. People over 65 were weighted 3 times the rest of the population to compensate for the 3x greater use of prescriptions in the elderly.

*Analysis*

Expression of the data year after year for each fundholding wave as an index with 100% equivalent to the continuing non-fundholders' cost and rate per age-weighted head of population - displayed as line graphs. Changes in cost, expressed as percentages, relative to changes in continuing non-fundholders displayed graphically year on year, beginning with the year before commencing fundholding.

The costs fell in each successive wave the year before fundholding by c.4%, this continued for the first year of fundholding, then began to rise again until by the 3rd year the original rate of change, similar to continuing non-fundholders (+/- 1%), was achieved. There was no perceptible change in rate of prescribing relative to continuing non-fundholders. Any overall reductions in cost (not quantified) were achieved by reduced cost per item rather than reduced rate of prescribing.

Mean relative prescribing cost reduction for fund-holders was less than 1.3%.

*Validity of conclusion*

The authors assert that the two greatest needs around fundholding are (a) to find a way of setting prescribing budgets that most effectively achieves reductions in costs and more importantly (b) an appropriate method of assessing their cost-effectiveness in terms of clinical outcome. Conclusion (a) does not follow; there is no evidence in the study that the budget-setting process itself achieved reductions, in fact all waves reduced costs in the year before starting, before any real budget was set. Conclusion (b) is supported by the identified reduction in cost per item, suggesting cheaper prescriptions, thereby raising the valid concern that reduced cost might be associated with reduced quality.

*Overall appraisal*

The waves of fundholders and the non-fundholders were different; despite aggregated data, they were all different in prescribing cost and rate prior to the study. Were the comparisons therefore valid ? - an issue recognised by the authors. The purpose of the study was achieved, but there is no estimation of

the statistical or fiscal significance of the changes. Clearly, whether significant or not, the changes had become similar to pre-fundholding in less than 3 years of being fundholders.

*Relevance to the thesis and main study*

Financial inducements (extrinsic motivators) are associated with transient effects (chapter 1). It is difficult to change prescribing rate (chapter 1). Studies of prescribing change utilising PPA data must be powerful enough to detect very small changes (of around 1% or so) as described in chapter 6, and supports the use of the number of prescriptions (a large number of events) as the denominator in analysis of the main study (chapter 7). Prescription items, while often criticised for being a crude way of measuring the volume of drugs prescribed, are nevertheless the best measure of the act of issuing prescriptions and therefore useful in assessing behaviour from a psychological point of view (chapter 5).

**Wilson RPH, Hatcher J, Barton S, Walley T. Influences of practice characteristics on prescribing in fundholding and non-fundholding general practices: an observational study. BMJ, 1996;313:595-9**

*Purpose of study*

To investigate (a) the influence of social deprivation, training status, partnership size and fundholding status on variation in prescribing between practices and (b) to investigate whether changes in prescribing after fundholding were attributable to fundholding alone or to the other factors.

*Population*

412 practices in the former Mersey region. Method of selection / exclusions were described elsewhere, but 384 practices ultimately took part.

*Method*

Cohort study with observations in year 1990/1 and repeated in 1993/4.

*Data*

Prescribing cost and rate (items), age-weighted population (over 65s multiplied by 3) from the NHS prescription database. The explanatory variable data was collected from health authority data and the low income scheme index (LISI) as an indicator of deprivation.

*Analysis*

Multiple regression analysis was applied to all waves of fundholding though differences were most marked between 1st wave fundholders and non-fundholders. Median cost per prescribing unit (age weighted individual) increased 28% less in the 1st wave fundholders relative to non-fundholders over the 4 year study period. In 1990/1 only 3% of the variation could be explained by the 'explanatory' variables. In 1993/4 only 7% of the variation could be explained by any model based upon the variables. Only 10% of the variation in changes between the observed years could be explained by the variables.

There was a 17% lower increase in prescribing rate in 1st wave fundholders relative to non-fundholders. In1990/1 34% of the variation in prescribing could be explained, and in1993/3 38%. The variable could not explain any of the variation in changes in rate between the observation years.

There was an 18% lower rise in cost per prescription item in 1st wave fundholders relative to non-fundholders. In 1990/1 43% of the variation, and in 1993/3 38% could be explained. Only 6% of the difference in changes in cost per item could be explained.

*Validity of conclusion*

The authors conclude that in neither year did fundholding make the major contribution to the variation in prescribing behaviour among practices (the other factors were more predictive), but it did seem largely responsible for differences in rise of costs between fundholders and non-fundholders. This is a valid summary.

*Overall appraisal*

The study was well executed, but the epidemiological approach could not establish causation, only association, between the variables and prescribing features of practices. Even the differences in change between groups of practices could have been due to features other than fundholding, or fundholding might be associated with unknown, perhaps unmeasurable, causative factors. The study aimed to investigate influences on prescribing, but really investigated the validity of proxy measures for the factors thought likely to have a bearing on prescribing cost and

rate. This information in itself may render resource allocation around prescribing more robust. The changes were quantified so it is possible for the reader to make a judgement on the significance of the effect of fundholding, bearing in mind that the groups used as a reference point, the non-fundholders, were different for many practice characteristics and were changing also. It is equally possible that, in an arena of rising cost and rate, that the fundholders represented the more stable group and that differences observed are due to de-motivated non-fundholders letting go of their hitherto tight control of prescribing. The study tackled the difficult topic of needs-led resource allocation in prescribing and raised some interesting questions.

*Relevance to the thesis and main study*

The explanatory variables were clearly associated with changes in prescribing rate and cost and validate the balance of these confounding factors (local social deprivation, size of partnership, training status and fundholding) across groups in the main study (chapter 6) and demonstrates the modest effect of fundholding (chapter 1).

**Eccles MP, Soutter J, Bateman DN, Campbell M, Smith JM. Influences on prescribing in non-fundholding general practices. BJGP, 1996;46:287-290**

*Purpose of study*

To establish (a) non-fundholding GPs' attitudes to a financial prescribing incentive scheme and (b) to document the presence and use of guidelines (prescribing policies) and (c) the influence of prescribing initiated in secondary care.

*Population*

All 448 non-fundholding practices in the Northern region.

*Method*

Postal questionnaire survey.

*Data*

Questionnaire addressed to first named principal for each practice and suggested completion by the most appropriate person in the practice. Questions covered influences on a practice's decision to attempt achievement of the incentive scheme target (including the influence of medical and

pharmaceutical advisers), presence and use of 3 named guidelines and perceived influence of hospital prescribing.

*Analysis*

The response rate was 78%. Practices most commonly perceived informal (45%) and formal (35%) discussions with partners as an important influence on whether to engage in the incentive scheme. Practices who achieved the target mostly saw the size of financial incentive as an important factor in their decision (57%), those who did not achieve felt most commonly that the target was unachievable (44%). Concerning the scheme, medical adviser input was found helpful by 67% of advised practices, and 71% of practices advised found pharmaceutical adviser input helpful.

Thirty-one percent of practices reported the possession of a prescribing policy or formulary and 85% said they always or usually used it. Practices who achieved their target were significantly more likely to possess guidelines. The most commonly possessed guidelines were asthma (89%), diabetes (87%) and hypertension (72%).

The most commonly reported areas where hospital prescribing was always felt to be influential was schizophrenia (30%), diabetes (16%) and ischaemic heart disease (10%).

*Validity of conclusion*

The authors' conclusion that GP prescribing is influenced by a complex web of factors with no factor pre-eminent seems vacuous. The questionnaire contained a complex web of questions and so gave rise to a complex web of answers. The impression is of 3 separate studies forced into one; study A explored attitudes toward a prescribing incentive scheme; study B attitudes and use of clinical guidelines and policies; and study C the effect of hospital initiated prescribing on GPs treatment decisions with regard to different diseases. Consequently a tangled web of conclusions emerged.

*Overall appraisal*

A good response rate to a comprehensive questionnaire addressing a host of treatment issues; the purpose was achieved. However, how much insight into influences on prescribing was gained seems obscure.

*Relevance to the thesis and main study*

Supports the validity of practice visiting by a medical adviser (chapter 6). When attempting to motivate practitioners they must perceive themselves capable of achieving the proposed task or no inducement will be sufficient (chapter 1).

## Bateman DN, Campbell M, Donaldson LJ, Roberts SJ, Smith JM. A prescribing incentive scheme for non-fundholding general practices: an observational study. BMJ, 1996;313:535-8

*Purpose of study*

To examine the effects of a financial incentive scheme on prescribing in non-fundholding general practices.

*Population*

459 non-fundholding practices in 1993/4 divided into three classes according to the ratio of their indicative prescribing amount to the local mean - A were = >10% above mean, B were 0% to 10% above mean, and C were below the mean.

*Method*

Observational study (with no control group) of the effect of the incentive scheme. The cost savings targets set for practices assumed that high cost practices could make savings more easily, therefore practices in group A were required to reduce cost by 3%, B by 2%, and C by 1%.

*Data*

Nationally held data on NHS prescriptions dispensed in 1993/4 practice by practice for non-fundholders. .

*Analysis*

Twenty-three percent of practices achieved their targets; 18% of practices in group A, 19% in group B and 27% in group C. Savings amounted to £1.54m at a cost of £463,000 in incentive payments. The region experienced the joint lowest prescribing over-spend in England in the study year. There was no overall significant difference in the rate of prescribing between achievers and non-achievers. Quality of prescribing scores (proportion of approved drugs prescribed) were higher in the achiever practices.

*Validity of conclusion*

The conclusion that practices in incentive schemes responded to financial incentives in a similar way to fundholders is not substantiated by the study, which did not compare non-fundholders with fundholders. The other conclusion that the incentive scheme did not seem to reduce the quality of prescribing is unsubstantiated, the achievers may have had even higher scores before intervention.

*Overall appraisal*

Recruitment of non-fundholding practices was total, which raises questions about the voluntary nature of the scheme. The absence of a control or reference group makes most of the measurements meaningless. The contribution of differences in historic under / over-funding between regions is not explored when implying that the scheme led to best prescribing performance in England. This seems to be an uncontrolled, unmatched, compulsory, unblinded trial based upon arbitrary prescribing targets.

Interestingly, a later study by the some of the same team (Roberts SJ, Bateman DN, Smith JM. Prescribing behaviour in general practice: the impact of promoting therapeutically

equivalent cheaper medicines. BJGP, 1997;47:13-17) shares similar methodological flaws.

*Relevance to the thesis and main study*

The study emphasises the scientific validity of the small matched controlled trial performed by this author and a colleague in Wirral in 1992/3 (chapter 2).

## Kelly MH, Murray TS. Motivation of general practitioners attending postgraduate education. BJGP, 1996;46:353-6

*Purpose of study*

To identify the motivating factors for GPs participating in continuing medical education before and after their new (1990) NHS contract.

*Population*

All 1959 GPs in the West of Scotland in the years 1990 and 1992.

*Method*

Questionnaire 1 (1990) was posted to all doctors and asked their views on continuing medical education.

Questionnaire 2 (1992) was a structured postal questionnaire sent to all GPs attending randomly selected courses inquiring as to what motivated the respondent to attend (i.e. the

factors causing attendance) and the reason for attendance (the justification for attending).

*Data*

The response rate in 1990 was 82.1% and 93.5% in 1992.

*Analysis*

The commonest prime motivating factors recorded in 1990 was interest (55.3%), improved knowledge (43.6%) and contact with colleagues (13.7%). In 1992 the commonest were interest (43.3%), postgraduate education allowance - PGEA - (34.6%) and improved knowledge (29.3%).

Reasons for attendance were recorded in 1992 the commonest was 'to gain sessions for PGEA'. Motivating factors for attending commercially sponsored meetings were significantly less for interest, significantly less for PGEA, significantly less for increased knowledge, and significantly more for venue/food/speaker.

*Validity of conclusion*

The authors conclude from the study that finance has a major effect on attendance at postgraduate meetings, but may not be a good incentive for learning. The first clause is valid, the

second somewhat speculative - learning was not measured although passive learning might be implied by the prominence of venue/food/speaker as a motivator to attend commercial meetings.

*Overall appraisal*

An interesting study that tends to confirm some intuitive fears about financial incentives and learning. The questionnaires were different and comparability is difficult to ascertain. The subjects were different in the 2 study years, 1990 were all GPs, 1992 were only attendees at PGEA meetings - are the two significantly different types of people. The results may not be generally applicable; the West of Scotland is a unique area with significant social, geographical and professional isolation in parts; the prominence of 'contact with colleagues' as a motivating factor perhaps bears this out.

*Relevance to the thesis and main study*

The paper confirms the inclusion of an interesting topic, improved knowledge and PGEA as motivating factors in the main study design (chapter 6) and also highlights concerns about the quality of commercially sponsored medical education (chapter 2).

## Clinical guidelines and best practice

### Newton J, Knight D, Woolhead G. General practitioners and clinical guidelines: a survey of knowledge, use and beliefs. BJGP, 1996;46:513-7

*Purpose of study*

To investigate (a) how familiar GPs are with a range of published guidelines, (b) to assess whether they have used them, (c) to describe their attitude to the guidelines, and (d) their attitudes the methods of implementing them.

*Population*

559 GPs randomly selected in the North and Yorkshire region.

*Method*

Postal questionnaire survey.

*Data*

Fifty-four percent response rate. Questions required opinions to be placed on a 5 point scale (positive = 5 in analysis)

*Analysis*

There was a higher proportion of single-handed GPs amongst the non-responders than responders. The BTS asthma guidelines were best known (mean score = 4.4), best used (3.8) and best changed practice (3.7), followed by the RCGP diabetes guidelines (3.3, 2.9, 3.4 and the RCR X-ray guidelines (2.9, 2.6, 3.1).

GPs' beliefs regarding guidelines were most strongly that they helped to improve knowledge (3.8), help you to learn more (3.7) and to use latest knowledge (3.5).

GPs felt most strongly that pressure to use guidelines emanated from the DoH (3.6), Medical Audit Advisory Groups (3.4) and FHSA managers (3.4).

Respondents felt most strongly that the implementation methods likely to have most impact were continuing medical education events (4.0), discussion with colleagues (4.0) and feedback on practice (3.9).

ANOVAs were performed on different sub-groups of GPs and showed that single-handers felt less strongly about the influence of discussion with colleagues and feedback on practice.

*Validity of conclusion*

The authors conclude that GPs are receptive to guideline initiatives, and their views are aligned with existing or proposed implementation strategies - this is consistent with the findings.

*Overall appraisal*

The response rate is low; non-responders may have held significantly different (possibly negative) views on the subject matter or indifference to it. Only 4 of the 13 questions regarding beliefs about guidelines were positively phrased, this could lead to biased responses. The knowledge claimed of the guidelines inquired about is likely to have been inflated (as recognised by the authors).

*Relevance to the thesis and main study*

Confirms that single-handed GPs are different and should be balanced across groups in the main study (chapter 6). Supports the use of an educational approach to guideline implementation

(chapter 3) and agrees that GPs are favourably disposed towards guidelines (chapter 4).

## Feder G, Griffiths C, Highton C, Eldridge S, Spence M, Southgate L. Do clinical guidelines introduced with practice based education improve care of asthmatic and diabetic patients? A randomised controlled trial in general practices in east London. BMJ, 1995;311:1473-8

*Purpose of study*

To determine whether locally developed guidelines on asthma and diabetes disseminated through practice based education improve quality of care in non-training, inner city general practices.

*Population*

Twenty-four self-selecting non-training practices in inner city east London. These were stratified for 6 relevant variables and randomly allocated to 2 groups, one to receive asthma guidelines, the other diabetes.

*Method*

Randomised control trial of practice based educational dissemination of guidelines. Balanced incomplete block design. Each of the two groups acted as the others' control. Education consisted of 3 lunchtime sessions at each practice with two educators. Session 1 introduced the guidelines and discussed how the practice's current management could be altered to suit. A stamp or booklet was given to act as a prompt during consultations and the session concluded with a discussion on the practicalities of home monitoring of peak flow or urine. Session 2 reviewed the practice's organisation and the clinical content of the guidelines. Session 3 took place 3 months later and reviewed performance and compliance with the guidelines.

*Data*

Prescribing data for the year before and year after the guidelines were introduced for asthma drugs were collected from the national NHS prescribing database. Quality of care scores were related to provision of specified diabetes relevant services.

*Analysis*

The diabetes guidelines practices improved significantly (analysis of covariance) in all 7 quality variables relative to control.

In the asthma practices improvement in asthma care was only significant in two variables. The prompts (rubber stamps and booklets) improved recording of variables for both conditions.

*Validity of conclusion*

The authors conclude that local guidelines disseminated via practice based education improve the management of diabetes and possibly asthma in inner city, non-training practices - this is consistent with the findings.

*Overall appraisal*

The sample practices were an unusual group, being inner city, non-training and electing to participate in a study; concerns are immediately raised about the universality of the findings. It seems a little unusual to stratify 24 subjects for six variables; as several variables are likely to coexist in one practice the randomisation process would have been interesting to observe. Minimisation would have been simpler and perhaps this is what, in effect, happened. The study is described as a randomised controlled trial, but the initial recruitment was open, only subsequent allocation to the experimental groups was randomised. The description of the educational meetings is not sufficient to judge whether the learning was passive or active.

Feedback, as in session 3, is a powerful confounding variable and may explain all the observed changes. The mixed result indicates that the hypothesis remains unproved.

*Relevance to the thesis and main study*

Two guidelines disseminated by practice based education had different effects, one was effective, the other not. Clearly some other variable is likely to be at work here; 'practice based education' itself is not sufficient otherwise both guideline disseminations would have been effective (chapter 3).

## Siriwardena AN. Clinical guidelines in primary care: a survey of general practitioners' attitudes and behaviour. BJGP, 1995;45:643-7

*Purpose of study*

To investigate GPs' (a) attitudes toward clinical guidelines, (b) their reported behaviour concerning them and (c) to investigate factors that might be associated with those attitudes and behaviour.

*Population*

All 326 GPs in Lincolnshire FHSA in February to October 1994.

*Method*

Postal structured questionnaire survey.

*Data*

Guidelines defined as. "statements designed to assist decision-making about appropriate care for a specified clinical condition". The questionnaire had sections on general details about the respondent, guidelines and audit. Negatively polarised attitudinal questions were reversed prior to analysis.

*Analysis*

The response rate was 65.3%, non-responders were similar in terms of age, sex and number of partners. Of these 78% had produced guidelines and 92% had performed clinical audit in their practice. GPs who had participated in audit were more likely to have produced guidelines, and members of the Royal College of GPs were more likely to have performed audit. The respondents felt most positively that good practice is not always scientific (4.04), they could exercise clinical judgement around guidelines

(3.91), and guidelines should be based upon what actually happens in practice (3.88). They felt most negative about not becoming a GP to practice 'cookbook medicine' (2.54), implementing guidelines demonstrates competence as a GP (2.55) and they would adopt clinical guidelines if there was a financial reward (2.72). In 9 of the 20 attitude questions RCGP members were significantly (paired t test) more positive than non-members.

*Validity of conclusion*

The conclusion that most GPs had produced in-house guidelines is technically correct, but the definition of 'clinical guideline' is so broad that it could encompass the most rudimentary statements of clinical policy. The positive attitude of RCGP members does indicate that the College could usefully take the lead in clinical guideline development in primary care.

*Overall appraisal*

The low response rate urges caution and the non-responders were unlikely to be similar in some respects not examined, for example membership of the RCGP, an important predictor variable, was likely to be under-represented in that group. Five point Likert scales with varied polarity are a good way to assess attitude. However, the resulting data is ordinal,

categorical, and not necessarily linear. In these circumstances median score rather than mean would have been more appropriate as would non-parametric analysis rather than the t test used.

*Relevance to the thesis and main study*

The attitude questionnaire is of similar design to the one I designed for assessing attitudes to clinical protocols (chapter 4). The generally positive disposition that GPs exhibit towards guidelines supports their incorporation in the study (chapter 6).

## Olesen F, Oestergaard I. Patients with urinary tract infection: proposed management strategies of general practitioners, microbiologists and urologists. BJGP, 1995;45:611-3

*Purpose of study*

To describe GPs', and specialists' (microbiologists' and urologists') strategies for diagnosis, treatment and follow-up of female patients with urinary tract infection symptoms.

*Population*

Two hundred GPs and 199 specialists (99 medical microbiologists and 100 urologists) were selected randomly from the register of the Danish doctors association. Twenty-two doctors were excluded for practical and eligibility reasons.

*Method*

Postal questionnaire using three vignettes to inquire into the respondent's clinical management in that situation; a girl aged 10 years, a 30 year old married woman, and a 60 year old woman - all previously healthy. For each fictitious case several diagnosis, treatment and follow-up options were offered.

*Data*

Response rates were; GPs 77%, microbiologists 51%, urologists 61%.

*Analysis*

Considerable variation was found within and between the groups of doctors. Specialists were more likely to deal with the 30 year old with telephone advice and to recommend a consultation for the 10 and 60 year olds. GPs were more likely to perform a urine culture test for the 30 year old, but specialists were more

likely to recommend antibiotic sensitivity testing. Both GPs and specialists were most likely to choose sulphamethoxazole as treatment, though specialists were more likely to recommend single-dose treatments. GPs suggested follow-up more than the specialists. All these differences were statistically significant (chi squared tests).

*Validity of conclusion*

The conclusion that there was large variation within and between the GP and specialist groups is substantiated. A second conclusion that small group based CME using specialists as a resource would not necessarily reduce variation in clinical practice between GPs seems bizarre, especially as the authors acknowledge the effectiveness of a CME campaign by microbiologists and GPs at promoting the use of sulphamethoxazole in Denmark. Another conclusion, that evidence based, rather than consensus directed guidelines would be needed in order to reduce variation in clinical practice is speculative and not supported by the study evidence.

*Overall appraisal*

A well-designed, ingenious questionnaire that was successful at evoking clinical decisions for three related conditions

that warranted differing management. Some features of the responses were unique to Denmark (e.g. the predominance of sulphamethoxazole treatment), this urges caution against universality of the findings. The analysis was simple and effective and the implications of the low response rates amongst specialists were explored. The authors raise alarm about the large within-group variation, but this was most prominent in the 30 year old case - a common condition without serious consequences, whether treated or not; in these circumstances some exercising of clinical freedom is not of great concern. Where good management is crucial, for example urine testing in childhood suspected urinary tract infection, there is consensus between the doctors and little variation.

*Relevance to the thesis and main study*

It would appear that both specialists and GPs would have a lot to gain by re-considering the management of urinary tract infection in the light of recent scientific evidence and to design up to date guidelines together - this supports the synthetic approach adopted in our preliminary study (chapter 4) and the choice of cystitis as the subject for the small clinical trial in the study (chapter 6).

## Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet, 1993;342:1317-22

*Purpose of study*

To review published evaluations of clinical guidelines that met defined criteria for scientific rigour.

*Scope*

Extensive search of electronic databases, bibliographies, reference trails and enquiries amongst academic colleagues to find published evaluations of clinical guidelines.

*Method*

Clinical guidelines were defined as systematically developed statements to assist practitioner decisions about appropriate health care for specific clinical circumstances. Only evaluative studies defined by the authors as rigorous were included; these were randomised controlled trials, cross-over trials, balanced incomplete block trials, controlled before and after studies and interrupted time series. The development, dissemination and implementation strategies for each were classed as 'internal' where it is implied that practitioners have been actively involved

(though this is not defined in the paper) and external where they appear to have been passively involved.

*Data*

Fifty-nine studies met the inclusion criteria.

*Analysis*

Following systematic review of each paper the distillation suggested that effective guidelines were associated with internal development strategies, dissemination through specific educational interventions and implementation through patient-specific reminders at time of consultation. Conversely, ineffective guidelines were associated with external / national development, dissemination by publication in journals and implementation through general reminders.

*Validity of conclusion*

The authors conclude that explicit guidelines do improve clinical practice when introduced in the context of rigorous evaluations, however, the size of the improvements in performance varied considerably. These are supported by the review findings.

*Overall appraisal*

A thorough and rigorous review of the published evidence. However, publication bias, the reluctance of authors to submit papers of ineffective initiatives and of editors to publish them, is likely to operate in any review of published evidence and was not considered by the authors. Contrasting and comparing the results of studies measuring the effect of differently developed guidelines by different variables, differently evaluated for different conditions in different health care systems is fraught with the danger of type 1 errors, and no statistical tool can help. Unless the reader carries out the same review he has to rely on the judgement of the reviewer as insufficient detail is available in the review paper to allow a valid appraisal. However, there is sufficient description of the method adopted to reassure the reader that as much care as reasonably possible has been taken by the authors and that their conclusions are worthy of consideration.

*Relevance to the thesis and main study*

The review supports the educational approach to the guideline initiative in the study (chapter 6). The 'internal' and 'external' classification is consistent with deep and superficial processing amongst recipients. The involvement of recipients in

the development of a guideline is associated with effectiveness as in our preliminary study (chapter 4).

# WIRRAL FAMILY HEALTH SERVICES AUTHORITY

## QUESTIONNAIRE

## REGARDING MEDICAL TREATMENT PROTOCOLS

Please circule the number along the scale that best represents your views. (Circling option (4) implies a neutral opinion).

**In the treatment of many medical conditions, are the public best served by :-**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| A diversity of approaches | (1) | (2) | (3) | (4) | (5) | (6) | (7) | A Uniform Approach |

**How often do you make use of currently available protocols ?**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Never | (1) | (2) | (3) | (4) | (5) | (6) | (7) | Always |

**Do you think that you will make more use of protocols in the future ?**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Definitely Yes | (1) | (2) | (3) | (4) | (5) | (6) | (7) | Definitely No |

**How does the availability of protocols make you feel ?**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Vulnerable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | Safe |

**Overall, do you think that the effect of protocols on medical practice is :-**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Beneficial | (1) | (2) | (3) | (4) | (5) | (6) | (7) | Harmful |

**At which level do you think that medical protocols should be developed ? (Please tick one box only)**

| | | | |
|---|---|---|---|
| National | ( ) | National and Local | ( ) |
| Local | ( ) | Local and Practice | ( ) |
| Practice | ( ) | National and Practice | ( ) |

All three ( )

**Any other comments ?**

_____

_____

_____

_____

Please return by FAX (or post) to Melanie Iredale, by 18th September, 1992. FAX No. 652 2668.

# DUODENAL ULCER MANAGEMENT PROTOCOL

**SOURCE: AXON (1991) BMJ: 919-920**

1 PATIENT presenting for the first time with symptoms:

Decide whether to endoscope or not.

- If yes and duodenal ulcer with H pylori is proven then treat with curative therapy

For 6-8 weeks
$H_2$ blocker OR colloidal bismuth OR sucralfate OR pirenzipine

Then follow with a course of triple H pylori eradication therapy for 2 weeks

| | |
|---|---|
| colloidal bismuthate subcitrate | 1 tab qds |
| + metronidazole | 400mg tds |
| + amoxicillin or tetracycline | 500mg qds |

- If no endoscopy, or if endoscopy does not prove H pylori then healing therapy only

For 6-8 weeks
$H_2$ blocker OR colloidal bismuth OR sucralfate OR pirenzipine

(But in selected patients a therapeutic trial of eradication therapy may be justified after a full explanation of the advantages and disadvantages).

---

2 PATIENTS with recurrent or resistant proven peptic ulcer disease.

An ulcer healing course of $H_2$ blocker OR colloidal bismuth OR sucralfate OR pirenzipine OR omeprazole

Then a course of triple H pylori eradication therapy as above.

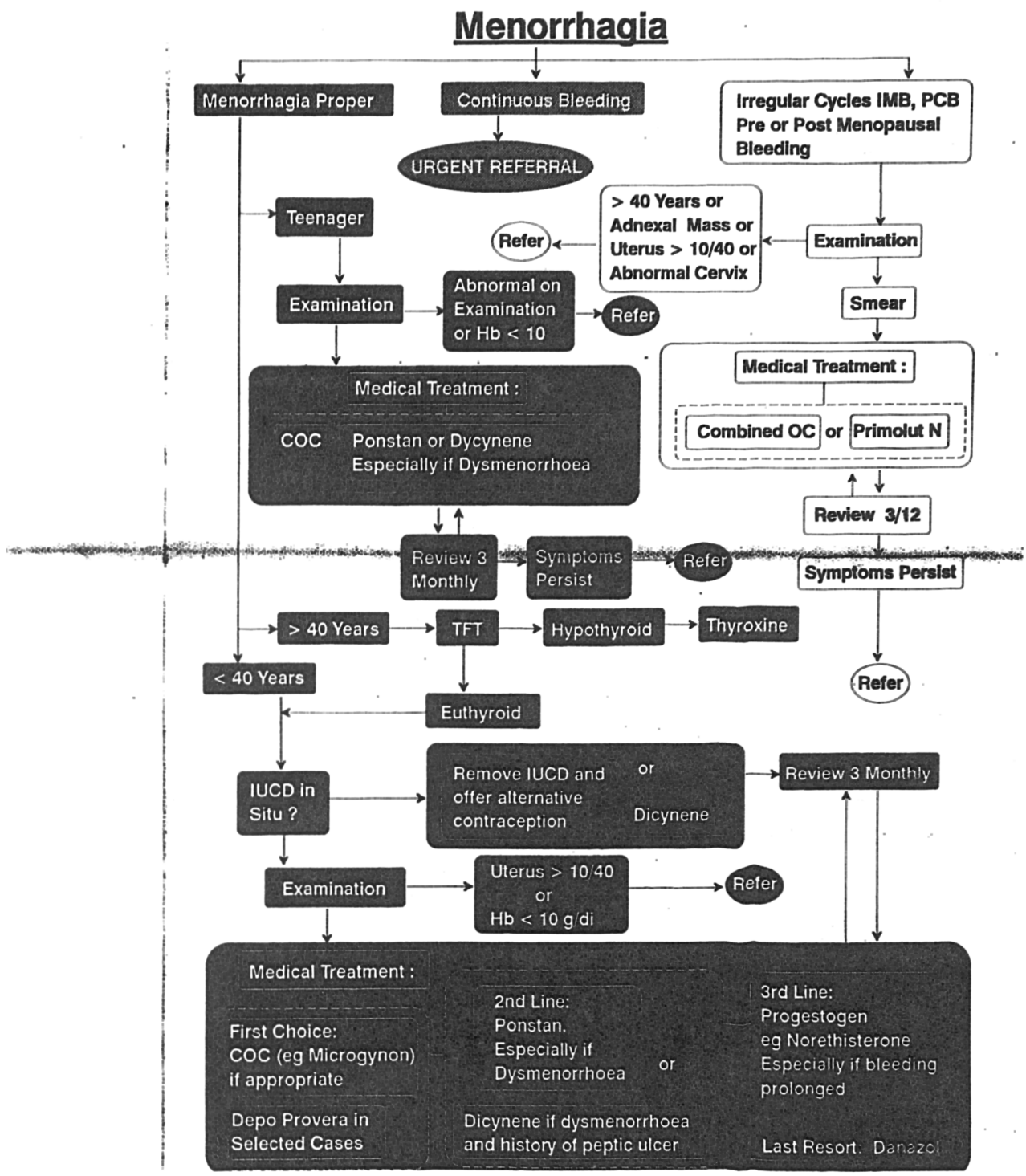
DOUBLE THERAPY
Bismuth for 8 weeks with metronidazole or tinidazole for the first 2 weeks will cause fewer side-effects, but is slightly less effective at eradicating H pylori. It may also be associated with antibiotic resistance and is not as well established or researched as triple therapy.

If duodenal ulcer disease continues despite H pylori eradication therapy, then maintenance $H_2$ blockers should be considered.

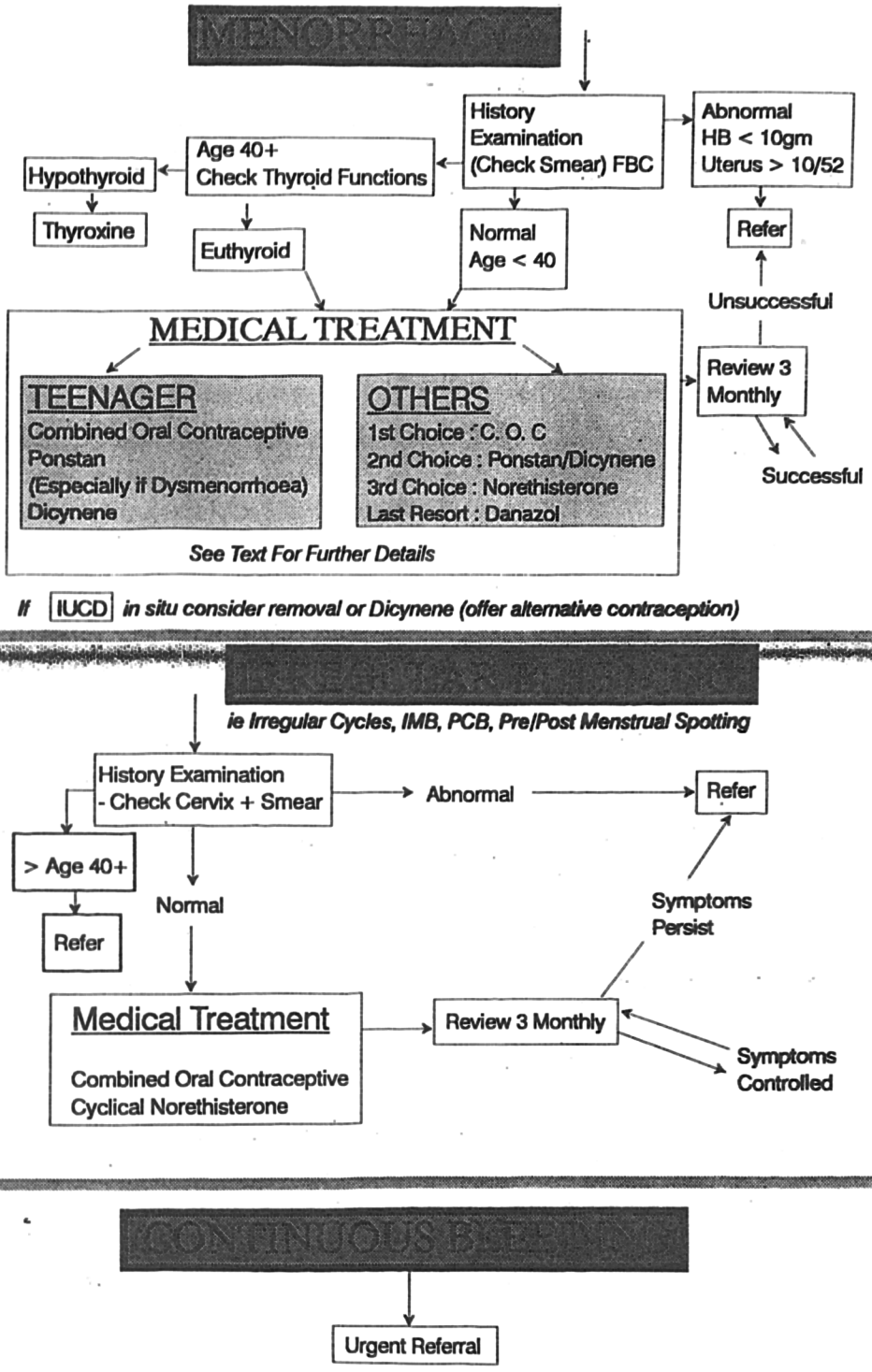
---

Source: Axon (1991) BMJ 919-920

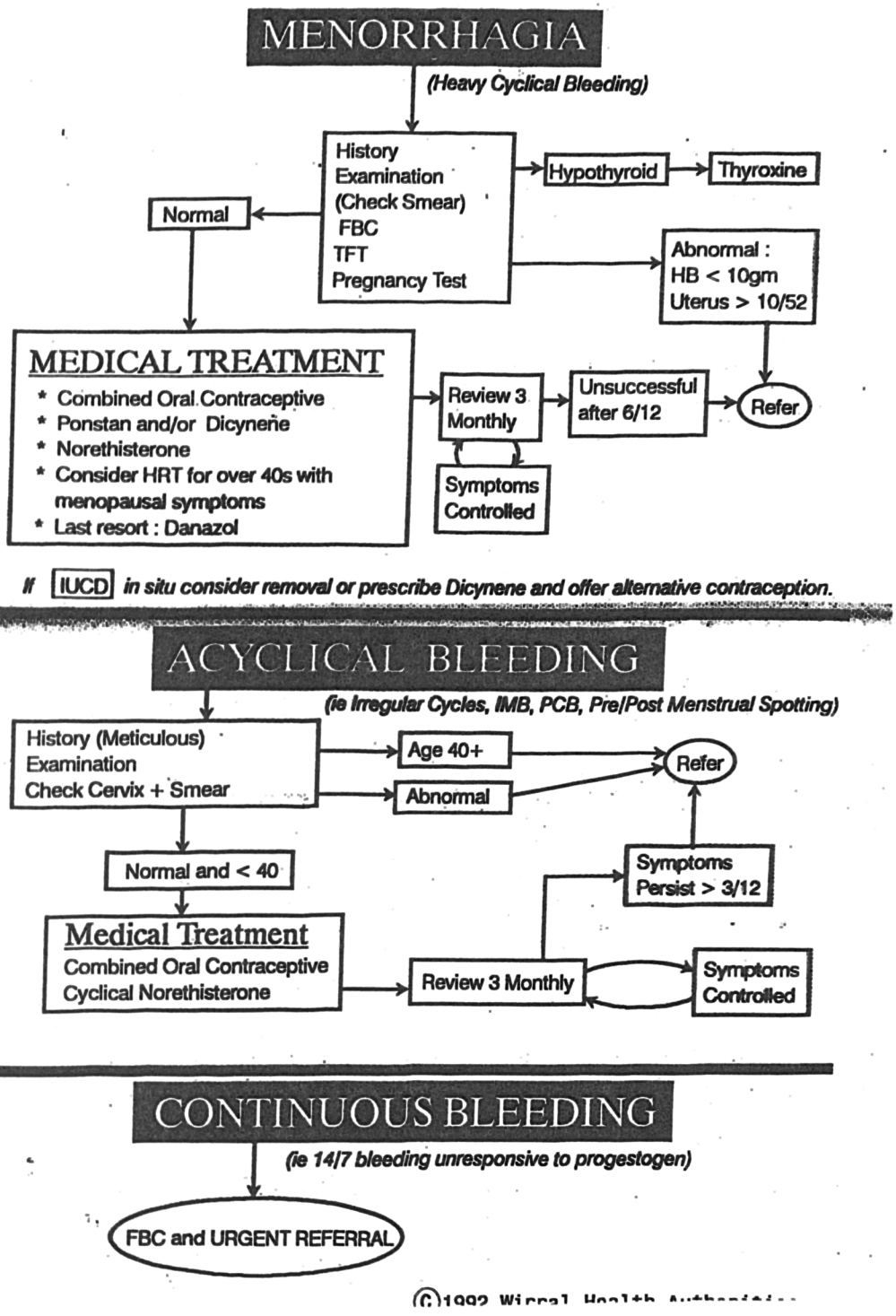
APPENDIX 4A
(Stage 1)



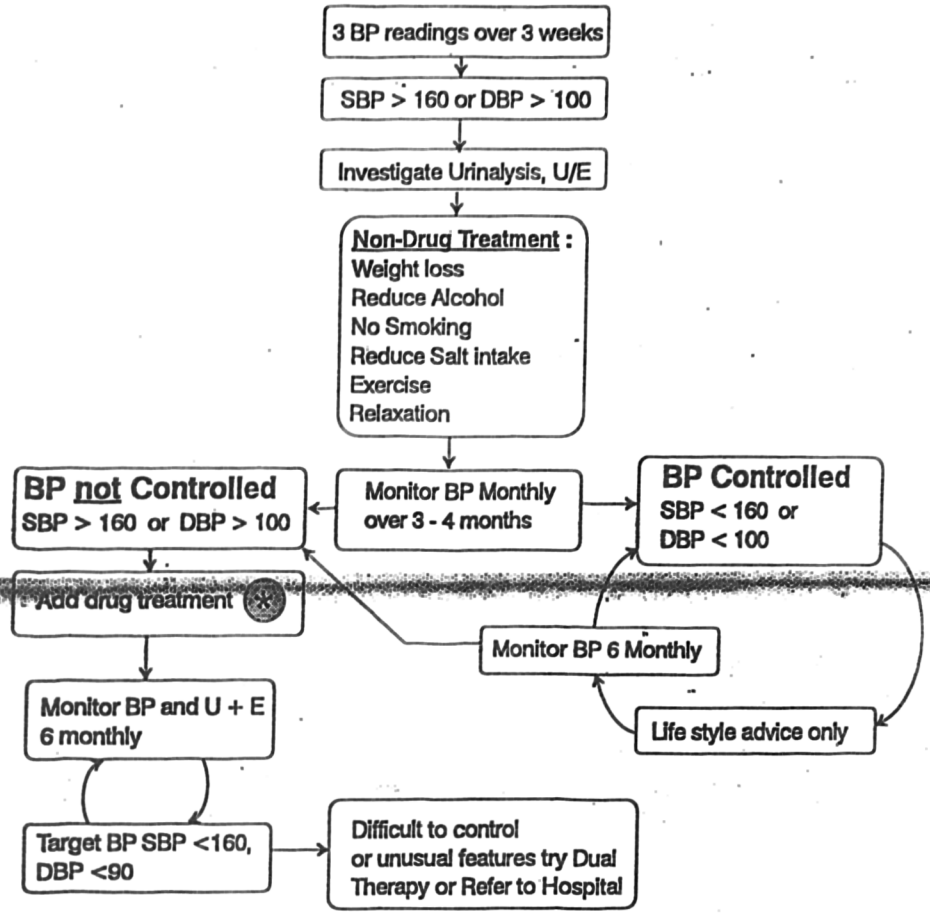
APPENDIX 4A
(Stage 2)



APPENDIX 4A
(Stage 3)

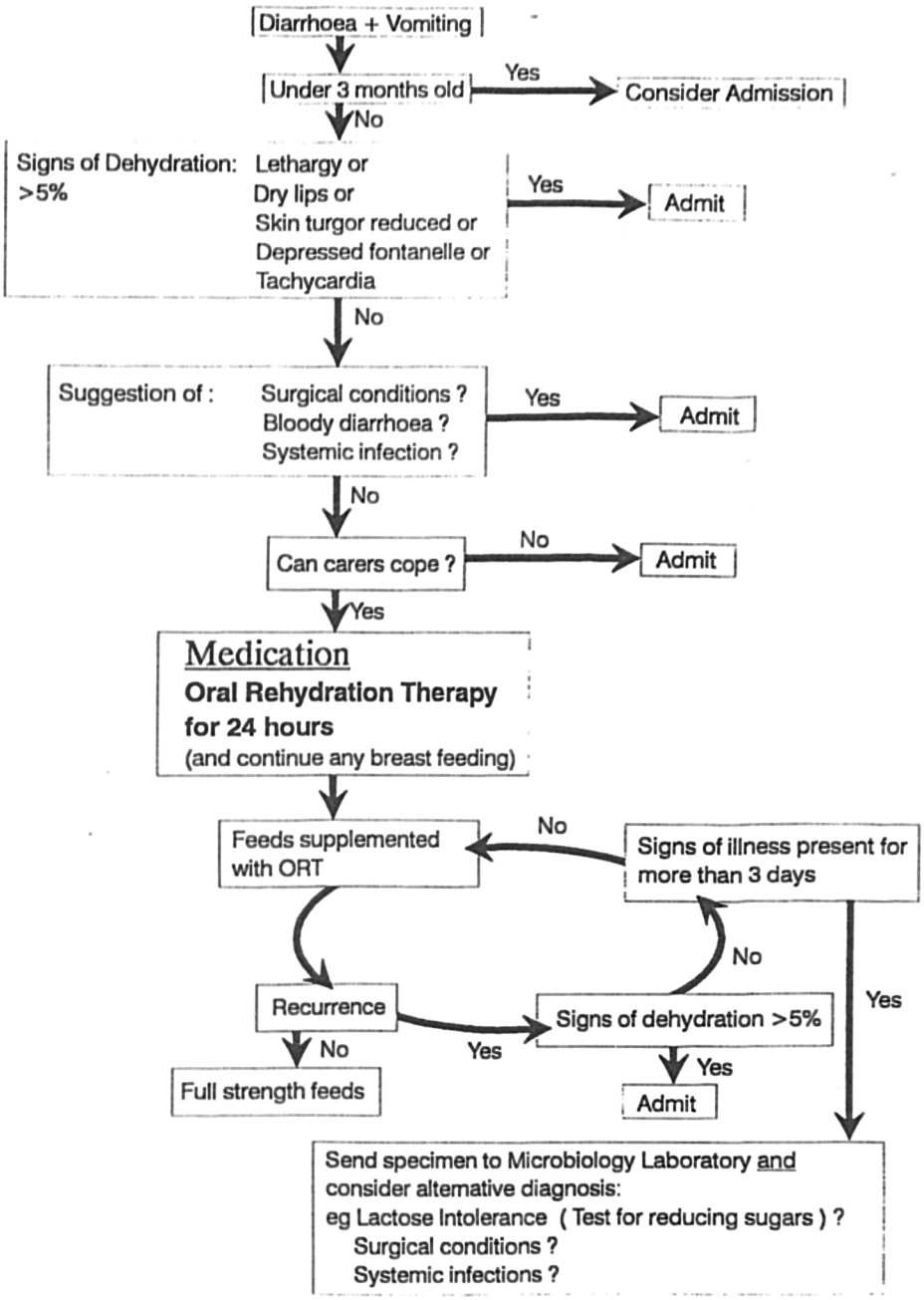


# HYPERTENSION IN THE ELDERLY




## DRUG TREATMENT

# Gastroenteritis in Children



APPENDIX 6A

Public Health Laboratory Service

Public Health Laboratory Service Board
61 Colindale Avenue
London NW9 5DF
Fax 081-200 8130/8131
Telephone 081-200 1295 - ext 3636

From The Director of the Service

Our ref Your ref

22 August 1994

Dear Dr Onion

GUIDELINES FOR THE MANAGEMENT OF INFECTION IN GENERAL PRACTICE

Thank you for your letter of the 14 July and for enclosing a copy of the above document. The work of Dr Bartzokas in Infection Control is well known to a number of my colleagues, both in Merseyside and of course here at Colindale. For example, we have read with interest the recent BMJ paper by Dr Bartzokas and colleagues on the importance of mouthflora to infections of prosthetic joints. It is also noteworthy that the production of this valuable set of guidelines has evolved from the collaboration between you and Dr Bartzokas.

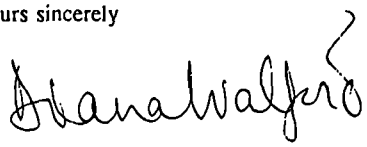
We are aware, of course, of the close working relationship between the PHLS laboratories at Chester and Liverpool in relation to the Wirral Health Authority now integrated with the Family Health Services Authority to form Wirral Health. We were also particularly glad to see the contribution that PHLS staff have made to 'Microbes' as indicated in the acknowledgements at the front of the publication.

The document incorporates a number of the principles and practices which of course are aspired to by other laboratories in the NHS and which the PHLS would certainly endorse as examples of good practice in its own laboratories throughout England and Wales. It is good to see that we are in complete harmony on so many of the points in your document. My colleagues and I found this a very comprehensive approach to the laboratory facilities as they apply to a modern microbiology service which incorporates analytical, clinical liaison, documentary support and management facilities under the leadership of the medical staff of Wirral Medical Microbiology.

I note that in your letter you indicate that the responses from GP users of the manual will be used to refine the document, and I would be grateful therefore if you could forward me a copy of the next edition following the returns from your GP colleagues. We also found

your application of Clinical Systems Analysis of interest and, again, if you detected evidence of persuading users to change their behaviour through the application of the advanced theories of social and cognitive psychology, my cõlleagues and I would be interested to hear of these developments.

Yours sincerely



DR DIANA WALFORD

Dr Carl Onion
Medical Director (Primary Care)
Wirral Health
St Catherine's Hospital
Clock Tower
Church Road
Tranmere
Birkenhead L42 OLQ

**Wirral general practitioner:**
**Preferred antibiotics questionnaire 1995.**

| Condition | Your antibiotic of first choice (before the results of any investigations) | Length of course (days) |
|---|---|---|
| URTI with catarrh | AMOXYCILLIN ✓ | 5 |
| Tonsillitis | Amoxil ✓ | 6 |
| Bronchitis | VELOSEF | 7 |
| Bronchopneumonia | VELOSEF | 7 |
| Lobar pneumonia | — VELOSEF | |
| Acute otitis media | AMOXYCILLIN ✓ | |
| Cystitis | TRIMETHOPRIM ✓ | 5 |
| Sinusitis | AMOXYCILLIN ✗ | 7 |
| Impetigo | AMOXYCILLIN ✗ | 5 |
| Cellulitis | AMOXYCILLIN ✗ | 7 |
| Gastroenteritis | CIPROXIN ✗ | 7 |
| Salpingo-oophoritis | — CIPROXIN ✗ | |

Any further comments you would like to make?

Thank you for taking the time to help us.

**Please fax to Dr C. Onion. Wirral Health on 652 2668**

**or post in the envelope provided.**

# Appendix PP

Published scientific papers arising from the thesis

Onion CWR, Walley T. Clinical guidelines: development, implementation and effectiveness. *Postgrad Med J*, 1995;**72**:3-9

Onion CWR, Slade PD. Depth of information processing and memory for medical facts. *Medical Teacher*, 1995;**17(3)**:307-313

Onion CWR, Dutton CE, Walley T, Turnbull CJ, Dunne WT, Buchan IE. Local clinical guidelines: description and evaluation of a participative method for their development and implementation. *Family Practice*. 1996;**13(1)**:28-34

Onion C. Rational antibiotic prescribing for LRTIs. *Prescriber*, 5th March 1996.

Onion CWR, Bartzokas CA. Imparting knowledge of best practice: deep versus superficial medical education. *British Journal of General Practice*. Submitted October 1996 - awaiting final confirmation.