

## **Evaluation of NHS Direct: impact and appropriateness**

An independent research study carried out by the Medical Care Research Unit of the University of Sheffield, on behalf of the Department of Health. The views expressed are those of the authors and not necessarily those of the Department of Health.

James Munro, Mike Clancy, Emma Knowles, Fiona Sampson, Jon Nicholl

with the assistance of Vicky Higham

October 2003

Medical Care Research Unit  
University of Sheffield  
Regent Court  
30 Regent Street  
Sheffield  
S1 4DA

© 2003 Medical Care Research Unit (MCRU), University of Sheffield

ISBN:

Published by the Medical Care Research Unit (MCRU), University of Sheffield

## Table of contents

1. EXECUTIVE SUMMARY .....	1
2. INTRODUCTION.....	2
3. RESEARCH AIMS .....	4
4. IMPACT OF NHS DIRECT: WAVE ANALYSIS.....	5
5. IMPACT OF NHS DIRECT: POPULATION SURVEYS .....	13
6. “APPROPRIATENESS” OF TRIAGE OUTCOMES .....	18
7. CONCLUSIONS.....	52
8. ACKNOWLEDGEMENTS .....	54
9. REFERENCES.....	55

## Index of tables

Table 4-1: Response rates to surveys of services in England and Wales	8
Table 4-2: Total calls to GP co-operatives	8
Table 4-3: Estimated annual % change in number of callers to GP co-ops, by wave	9
Table 4-4: Total first attendances at A&E departments	10
Table 4-5: Estimated annual % change in number of first attendances at A&E, by wave	10
Table 4-6: Total 999 journeys by ambulance services (by health authority served)	11
Table 4-7: Estimated annual % change in number of 999 journeys by ambulance services, by wave	11
Table 5-1: Respondents to the population survey	14
Table 5-2: Health problem in the previous four weeks for which help or advice was sought*	14
Table 5-3: Changes in service use for unplanned health events	15
Table 5-4: Use of NHS Direct for unplanned health events	16
Table 6-1: Initial events included in the postal survey	25
Table 6-2: Additional events derived from free text data	26
Table 6-3: Drug events derived from free text data	28
Table 6-4: Selected immediate care services available in study areas	30
Table 6-5: Age and gender of survey respondents	34
Table 6-6: NHS Direct triage advice reported by patient	36
Table 6-7: NHS Direct triage advice, after recoding	37
Table 6-8: How soon NHS Direct advised contacting another service, reported by patient	37
Table 6-9: Whether the patient acted on the advice (Method A: self-report)	38
Table 6-10: Whether the patient acted on the advice (Method B: derived)	38
Table 6-12: How easy it was to follow the advice in practice	39
Table 6-13: Patients' views of appropriateness of advice	39
Table 6-14: Patient's views of appropriateness of timing advised	39
Table 6-15: Patients' views of helpfulness of advice	40
Table 6-16: Contacts with services after the call to NHS Direct	40
Table 6-17: Necessariness of first action taken, assuming full service availability	42
Table 6-18: Necessariness of first action taken, varying assumption of service availability (site 2 only)	43
Table 6-20: Necessariness of first contact, compliers only (Method B)	44
Table 6-21: Necessariness of first contact by patient's view of appropriateness of advice given	47
Table 6-22: Necessariness of first contact by patient's view of appropriateness of advice given, compliers only (Method A)	48

## 1. EXECUTIVE SUMMARY

NHS Direct, the national nurse-led telephone helpline, was announced in the white paper *The New NHS: modern, dependable* in December 1997, and has been available to the population of England and Wales since November 2000, and in parts of Scotland (as NHS24) since mid-2002. The purpose of the service is to provide “easier and faster advice and information for people about health, illness and the NHS so that they are better able to care for themselves and their families”.

A number of evaluations of NHS Direct have been published, including three progress reports from the Medical Care Research Unit, which have addressed issues of activity, costs, casemix, callers' views, compliance, nurse and other stakeholder views, impact and safety. The aims of the work reported here were to determine the appropriateness of actions and impact on services attributable to NHS Direct, for a further period and in a wider range of sites than previously examined.

Using data from ambulance, A&E and GP co-operative services across England, Wales and Scotland, and examining the effect of the introduction of each successive wave of NHS Direct, we conclude that NHS Direct has been associated with a fall in the number of calls to co-operatives, but with no important change in demand for A&E or emergency ambulance services. Using large population surveys in the areas longest served by NHS Direct, over four successive years, we draw similar conclusions. These data suggest no important change in the use of other services, and no change in the mean number of health care services used during an unplanned episode of care. They also suggest that, although NHS Direct call rates are rising, the service is used in only a small proportion (6% or less) of unplanned episodes of care.

We have examined the appropriateness of NHS Direct triage advice in 2748 calls in terms of the “necessariness” of contacts with services which follow the call. Under the most realistic assumption, that nurse-led services (minor injury units and walk in centres) are unavailable locally, this study suggests that advice leads to necessary contacts in 86% of calls, unnecessary contacts in 13% of calls, and insufficient contacts in about 1% of calls. Further work is underway to determine whether the reasons for apparent over-triage are evident during the call.

The process-based method of call assessment developed and used during this study has the potential to be more widely applied within NHS Direct to evaluate the effects of service improvements, and as a tool for clinical audit.

The results we have reported here suggest that NHS Direct triages to necessary and sufficient services in a large majority of cases. There is also evidence, however, that at least one in eight callers may receive advice leading to an unnecessary contact with health services. The challenge for NHS Direct now is to reduce this proportion, while at the same time maintaining the safety of the service.

## 2. INTRODUCTION

### 2.1 Policy background

NHS Direct, the national nurse-led telephone helpline, was announced in the white paper *The New NHS: modern, dependable* in December 1997,<sup>1</sup> following one of the recommendations made in the Chief Medical Officer's report *Developing Emergency Services in the Community*, published 3 months earlier.<sup>2</sup> The white paper indicated that the purpose of the new service would be to provide "easier and faster advice and information for people about health, illness and the NHS so that they are better able to care for themselves and their families". In addition, the Chief Medical Officer's report had expressed the hope that such a service might help to "reduce or limit the demand" on existing immediate care services.

The specific objectives set by the Department of Health for NHS Direct include:

- to offer the public a confidential, reliable and consistent source of professional advice on healthcare, 24 hours a day, so that they can manage many of their problems at home or know where to turn to for appropriate care;
- to provide simple and speedy access to a comprehensive and up to date range of health and related information;
- to help improve quality, increase cost-effectiveness and reduce unnecessary demands on other NHS services by providing a more appropriate response to the needs of the public; and
- to allow professionals to develop their role in enabling patients to be partners in self-care, and help them to focus on those patients for whom their skills are most needed.

The development of NHS Direct since the initial policy commitment was rapid, and by November 2000 the telephone service was available to the entire population of England and Wales. A similar service for Scotland, NHS 24, was announced in December 2000 and began taking calls in the Grampian region in mid-2002.<sup>3</sup>

### 2.2 Evaluation to date

A number of evaluations of NHS Direct have been published, including three progress reports from the Medical Care Research Unit.<sup>4 5 6</sup> Broadly, this research has shown that NHS Direct is a popular and well-used service which callers find friendly, helpful, prompt and reassuring.<sup>7</sup> Population call rates are at about one-third the level of out-of-hours general practice, and the pattern of use is characterised predominantly as an out-of-hours service, with relatively high call rates for young children and young women, and lower than expected call rates for older adults. The great majority of callers are seeking advice on the immediate management of a wide range of mainly minor illnesses and injuries.

However, not all callers have been happy with the service. Important sources of dissatisfaction have included difficulties in getting through to the service or delays in being able to speak to a nurse,<sup>8</sup> or the number of questions asked.<sup>4</sup> Anecdotal media reports,<sup>9</sup> and evidence from the Consumers' Association,<sup>10</sup> have raised questions about the quality and consistency of self-care advice, although the methodology of the Consumers' Association study has been criticised for providing no comparison with the quality of advice offered by other health care professionals.<sup>11</sup> Nonetheless, there is currently little or no evidence available on the quality of self-care advice.

There is some evidence to suggest that the introduction of NHS Direct was associated with halting an upward trend in demand for out-of-hours general practice, but has had no measurable impact on overall demand for ambulance or accident and emergency services,<sup>12</sup> or on winter GP consultations for flu-like illness.<sup>13</sup> The evidence available also suggests that serious adverse clinical outcomes associated with NHS Direct are likely to be rare.<sup>6</sup>

### **3. RESEARCH AIMS**

The aims of the proposed research were to determine the appropriateness of actions and impact on services attributable to NHS Direct, for a further period and in a wider range of sites than previously examined, and during a period in which call volumes were increasing and a new clinical assessment system had been introduced. The overall purpose was to contribute to the further development of NHS Direct.

#### **3.1 Objectives**

The objectives of the research were:

- a) To assess the impact of successive national waves of NHS Direct on demand for ambulance, A&E and GP out-of-hours services;
- b) To continue to assess the impact of NHS Direct on the overall volume and distribution of demand for health care;
- c) To assess knowledge of availability of new health services in the general population;
- d) To determine the proportion of contacts with immediate care services, following advice from NHS Direct, which prove to be necessary;
- e) To identify, and list the reasons for any inappropriate contacts, which may include inappropriate triage advice from NHS Direct.
- f) To identify, and list the reasons for any inappropriate advice which may have been given.
- g) To determine callers' views on the appropriateness and helpfulness of the advice they receive.

Ethical approval for the study was granted by Trent Multi-centre Research Ethics Committee.

## **4. IMPACT OF NHS DIRECT: WAVE ANALYSIS**

### **4.1 Introduction**

We have previously examined the effect of the introduction of NHS Direct in the three first wave areas on the use of immediate care services in these areas.<sup>12</sup> While there was an association with a “levelling off” in the increasing demand for out-of-hours general practice, there was no measurable effect on use of emergency ambulances or attendance at A&E departments.

As time has passed, NHS Direct has covered progressively larger geographical areas and has handled increasing numbers of calls. In addition, a new decision support system was introduced during 2001 which may have changed the distribution of triage decisions.<sup>14</sup> Such developments mean that the “dose” of NHS Direct applied to the population may now be much greater than it was in its first year, so that an effect on trends in demand for other services may now be apparent.

However, the fact of national (England and Wales) coverage means that there are now no control areas untouched by NHS Direct with which those areas served can be compared. In order to overcome this problem, we undertook a “wave” analysis in which any change in trend in those new areas served by each additional wave of NHS Direct is compared with the change in all other areas not changing at that time.

### **4.2 Methods**

#### **4.2.1 Data sources**

The data sources for this analysis were the routine activity records maintained by A&E departments, GP co-operatives and ambulance services. We carried out a postal survey of all such services in England, Wales and Scotland asking for monthly figures on the number of first attenders (A&E departments), calls from patients (GP co-operatives) and emergency ambulance journeys (ambulance services). In the case of ambulance services, we asked for the data to be provided separately for each health authority which the ambulance service covered. The time period for which data was sought was April 1997 (one year before NHS Direct began) to March 2001.

In addition, we asked each service for the date when NHS Direct covered at least half of their catchment population, in order to assist in allocating each service to an NHS Direct wave. Further data on when each service was covered by NHS Direct was obtained from Department of Health policy documents and checked with local NHS Direct services.

Non-respondents were sent up to two reminders and in many cases telephoned in order to maximise the response rate.

For the analysis, we assumed that NHS Direct was introduced in four discrete waves: wave one in March 1998, wave two in March 1999, wave three in December 1999 and

wave four in November 2000. Each responding service was allocated to a single NHS Direct wave. The allocation of waves was checked by reference to press releases and Department of Health documents, and by consulting each NHS Direct site to resolve questions about coverage at each stage of expansion. Services in Scotland which provided data were allocated to “wave five”, i.e. not covered by NHS Direct or NHS24 throughout the period of the study. This data was included in the analysis, providing “pre-NHS Direct” data in each time period.

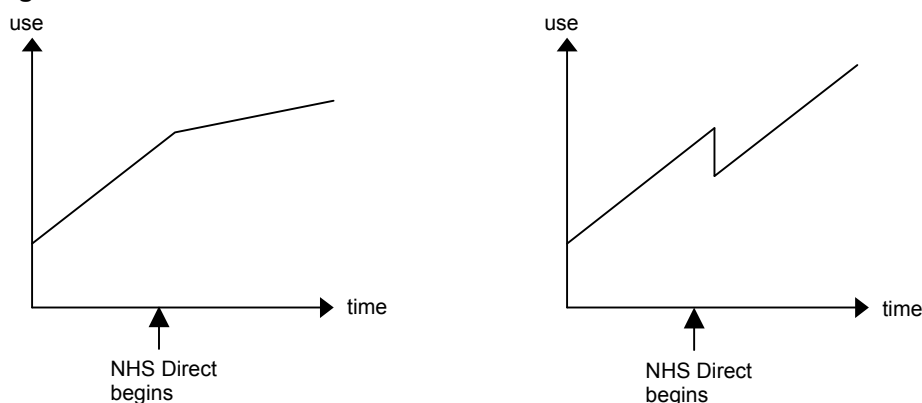
#### 4.2.2 Data analysis

The effect of NHS Direct on the use of services was examined by fitting statistical models to the monthly numbers of emergency ambulance journeys, A&E attendances, and calls to GP co-operatives. The purpose of each model was to remove the effects of long term trends in service use and of any systematic monthly or seasonal variation. The models were not time-series models in the conventional sense in which allowance is made for correlation between successive observations (autocorrelation). Rather, the models used here treated the effects as simple fixed effects, in order to remove any systematic variation in service use before estimating and testing whether the introduction of NHS Direct was associated with any change in service use.

In principle, NHS Direct might have one or both of two distinct effects. First, it might introduce a change in the long term *trend* in service use, for example turning an upward trend in the use of co-operatives into a downward trend. This idea is illustrated in the left hand diagram in Figure 4-1, below. Second, it might also be argued that NHS Direct would have a single, one-off *step* effect, removing a fixed proportion of existing service use, but have no effect on the underlying trend in use. Of course, it is also plausible that both effects might operate.

There are reasonable arguments in favour of either model, and no strong reason to prefer one over the other. Therefore, in the analysis reported here, we have separately applied each of these models to the available data.

**Figure 4-1: Possible effects on NHS Direct on service use**



The models were fitted to the log(counts) of journeys, attendances and calls, for the 48 months from April 1997 to March 2001 spanning the introduction of NHS Direct.

In the first (“trend”) approach, the effect of NHS Direct was estimated by first fitting a linear time trend to the monthly counts, and then adding a further linear time trend which started when the co-op, A&E department, or health authority catchment area was first covered by NHS Direct. Since NHS Direct was introduced in four waves this additional trend could start at one of four different times ( $t = 12, 24, 33, 44$  months), depending on where the particular service was located. However, for each type of service (co-ops, A&E, ambulances) the regression coefficient for the “post-NHS Direct effect” was constrained to be the same in all four waves so that it estimated the *same* effect of NHS Direct, relative to the underlying trend, for *all* waves.

In the second (“step”) approach, an initial linear trend was fitted to monthly counts in the same way as before, followed by a term for the step change in use when the service was first covered by NHS Direct. Again, the proportionate step change was constrained to be the same for all four waves

Models which assumed Poisson-distributed counts and were fitted to the square root of the counts were also examined. These gave essentially the same results and only the multiplicative models are reported here. Models with quadratic terms in the underlying trend, or the trend effect of NHS Direct, rarely improved the fit and a simple linear model was used throughout. None of the models showed a strong empirical basis for preferring either the trend or the step analysis.

### **4.3 Results**

#### **4.3.1 Data completeness**

The number of services responding to the request for data, and the number providing data which could be used in the analysis, are shown in Table 4-1 below.

“Inappropriate respondents” were units which did not have a minor injury or A&E service, or whose data had been already been returned as part of another service.

Organisations unable to supply data included those which supplied less than 12 months data, supplied data annually or quarterly, or could not supply data at all.

Those with missing data included those with a data series which began later than April 1997, stopped earlier than March 2001, or contained gaps in the series. One A&E department could not be reliably allocated to a specific NHS Direct wave.

There was no ambulance service data for first wave sites in Milton Keynes and Preston, Chorley and South Ribble, since it could not be separated out from the rest of the data for the surrounding health authority. The data for these entire health authorities was therefore classified as incomplete because it included data from more than one wave.

**Table 4-1: Response rates to surveys of services in England and Wales**

	A&E departments	GP co-operatives	Ambulance services	Health authorities (from ambulance service data)
Number surveyed	239	297	35	109
No response	30	101	0	0
Inappropriate respondent	9	8	0	0
Unable to supply data	14	34	2	-
Missing data	15	50	-	25
Uncertain wave	1	0	-	9
<b>Usable in analysis</b>	<b>170</b>	<b>104</b>	<b>-</b>	<b>75</b>

#### 4.3.2 GP co-operatives

The number of co-operatives in each wave, and the total number of calls made to co-operatives available for the analysis, by wave and time period, are shown in Table 4-2 below.

In the first (“trend”) approach to analysis, there was a moderate and statistically significant fall in the underlying trend in GP co-op call numbers with the introduction of NHS Direct (Table 4-3, in which the shaded area denotes the waves covered by NHS Direct in any given time period). The underlying annual trend fell from -0.1% without NHS Direct to -6.8% with NHS Direct, a reduction of -6.7% (95% CI -7.5% to -5.9%).

The total number of calls received each month, adjusted for the seasonal month effects, showed a substantial increase in the number of calls during the first 11 months of our study period (April 1997 to January 1998) before first wave sites started, followed by a small decline over the next four study periods during which the four waves were introduced (Figure 4-2). However, the estimated annual percentage change in each study period was similar for co-ops in each wave (Table 4-3) and hence the change from period one to subsequent periods cannot have been due to NHS Direct, which by definition did not cover co-ops in these waves until later.

Nevertheless, omitting the first 11 months and fitting the model only to months 12 to 48 hardly changed the estimated effect of NHS Direct. For the period March 1998 to March 2001 only, NHS Direct was estimated to have reduced the monthly number of calls by -5.0% (95% CI -4.3% to -5.8%) per annum.

In the second (“step”) approach, there was a one-off drop in use of co-operatives by 8.6% (95% CI -9.8 to -7.5). Given the mean demand on the GP co-operatives in this study of 94 calls per co-operative per night, this is equivalent to a fall of 8 calls per night for the “average co-op”.

**Table 4-2: Total calls to GP co-operatives**

Wave	No of co-ops	No of calls	No of calls			
			Apr 97-Feb 98	Mar 98 - Feb 99	Mar 99 - Nov 99	Dec 99 - Oct 00 Nov 00 - Mar 01

1*	-	-	-	-	-	-	-
2	28	5230048	1107082	1361233	946167	1234089	581477
3	15	1811204	383778	461240	320350	436415	209421
4	35	4584925	939195	1166117	816080	1132850	530683
5	25	2533575	550103	636991	450394	608399	287688
<b>All</b>	<b>103</b>	<b>14159752</b>	<b>2980158</b>	<b>3625581</b>	<b>2532991</b>	<b>3411753</b>	<b>1609269</b>

\* There were no wave 1 co-ops returning usable data over the four year period.

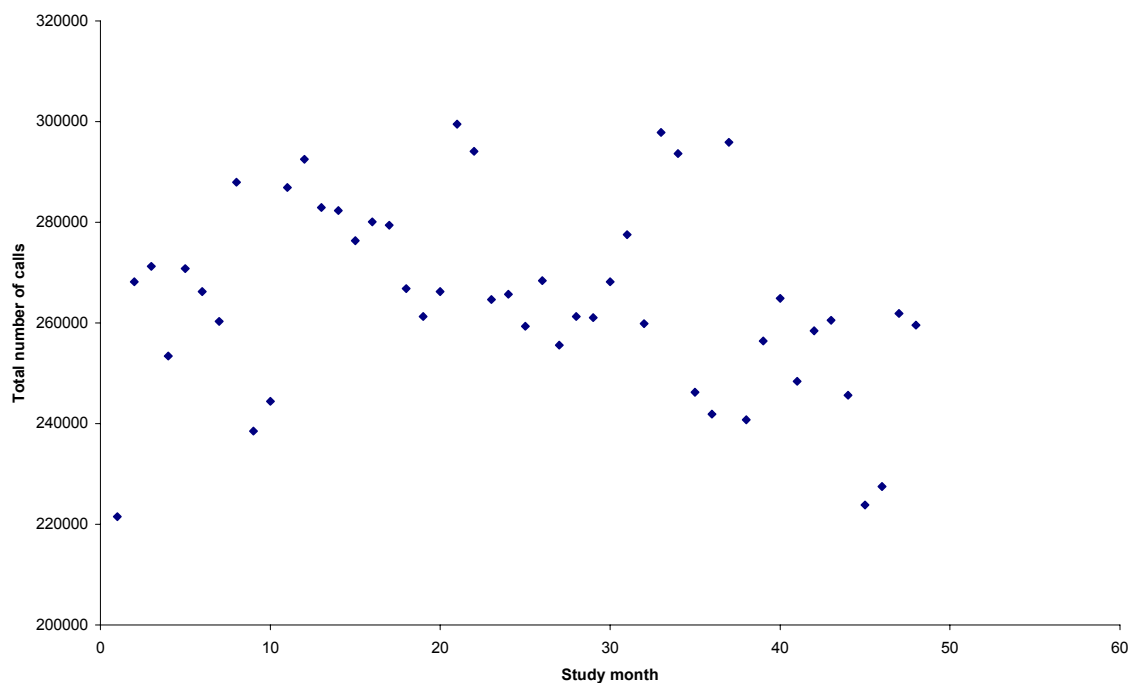
**Table 4-3: Estimated annual % change in number of callers to GP co-ops, by wave**

Wave	No of co-ops	Apr 97 to Feb 98	Mar 98 to Feb 99	Mar 99 to Nov 99	Dec 99 to Oct 00	Nov 00 to Mar 01
1	-	-	-	-	-	-
2	28	+16.6	-5.5	-6.1	-13.5	+9.1
3	15	+13.0	-3.0	-1.6	-9.3	+2.4
4	35	+16.4	-0.1	+1.4	-5.8	-0.0
5*	25	+12.6	-5.3	+4.6	-4.6	-6.0
<b>All</b>	<b>103</b>	<b>+15.0</b>	<b>-3.3</b>	<b>-0.3</b>	<b>-8.2</b>	<b>+1.2</b>

These figures denote the estimated mean linear trend in each wave and period.

\* Wave 5 comprises control co-ops which did not introduce NHS Direct until after March 2001.

**Figure 4-2: Total number of calls to GP co-ops by study month (April 97 to March 2001) adjusted for systematic monthly variation**



#### 4.3.2.1 A&E departments

The number of A&E departments in each wave, and the total number of first attendances available for the analysis, by wave and time period, are shown in Table 4-4 below.

In the trend analysis, there was weak evidence of a very small increase in the underlying upward trend in number of first attendances at A&E departments after the introduction of NHS Direct, relative to controls. The annual trend was estimated to have increased by +0.4% per year (95% CI 0.1% to 0.7%), from +0.91% per year to +1.3% per year (Table 4-5).

In the step analysis, there was a one-off step up in use by 1.2% (95% CI +0.6 to +1.8). Given that the A&E departments in our sample saw an average of 146 patients per department per day, this is equivalent to an increase of between one and two patients per day for the “average department”.

**Table 4-4: Total first attendances at A&E departments**

Wave	No of depts	No of attendances	Apr 97-Feb 98	Mar 98 - Feb 99	Mar 99 - Nov 99	Dec 99 - Oct 00	Nov 00 - Mar 01
1	4	1235155	282935	308635	242446	281032	120107
2	61	14801265	3330680	3631824	2865383	3471176	1502202
3	33	7114600	1613183	1774616	1375045	1641955	709801
4	69	12365261	2816443	3049983	2395110	2874516	1229209
5*	-	-	-	-	-	-	-
<b>All</b>	<b>167</b>	<b>35516281</b>	<b>8043241</b>	<b>8765058</b>	<b>6877984</b>	<b>8268679</b>	<b>3561319</b>

\* There were no wave 5 A&E departments with complete data in the study

**Table 4-5: Estimated annual % change in number of first attendances at A&E, by wave**

Wave	Apr 97 to Feb 98	Mar 98 to Feb 99	Mar 99 to Nov 99	Dec 99 to Oct 00	Nov 00 to Mar 01
1	-0.42	-0.2	+7.5	-10.2	+10.1
2	-1.9	+2.0	+6.8	-2.2	+0.7
3	+0.7	+0.6	+3.9	-2.8	+1.1
4	-1.5	-0.3	+5.6	-2.6	+1.4
5	-	-	-	-	-
<b>All</b>	<b>-1.2</b>	<b>+0.7</b>	<b>+5.7</b>	<b>-2.7</b>	<b>+1.3</b>

These figures denote the estimated mean linear trend in each wave and period.

#### 4.3.2.2 Ambulance services

The number health authority areas in each wave, and the total number of 999 ambulance journeys available for the analysis, by wave and time period, are shown in Table 4-6. In the step analysis, there was a one-off step *down* in use by 1.2% (95% CI –2.0 to –0.4). Given that the ambulance services in our sample undertook an average of 217 journeys per service (not per health authority) per day, this is equivalent to a fall of between two and three journeys per day for the “average service”.

Table 4-6.

In the trend analysis, the underlying trend of +4.0% increase per year in ambulance service use was estimated to have reduced slightly, by –0.4% (95% CI –0.8% to +0.1%), with the introduction of NHS Direct (Table 4-7).

In the step analysis, there was a one-off step *down* in use by 1.2% (95% CI –2.0 to –0.4). Given that the ambulance services in our sample undertook an average of 217 journeys per service (not per health authority) per day, this is equivalent to a fall of between two and three journeys per day for the “average service”.

**Table 4-6: Total 999 journeys by ambulance services (by health authority served)**

Wave	No of HAs	No of journeys	Apr 97-Feb 98	Mar 98 - Feb 99	Mar 99 - Nov 99	Dec 99 - Oct 00	Nov 00 - Mar 01
1	2	223711	48618	54737	41508	53507	25341
2	23	4000381	863576	985069	749767	957319	444650
3	12	1612881	348243	399082	303109	383346	179101
4	22	3217246	688637	794215	610678	770836	352880
5*	16	1392813	291260	343401	264511	337694	155947
<b>All</b>	<b>75</b>	<b>10447032</b>	<b>2240334</b>	<b>2576504</b>	<b>1969573</b>	<b>2502702</b>	<b>1157919</b>

**Table 4-7: Estimated annual % change in number of 999 journeys by ambulance services, by wave**

Wave	Apr 97 to Feb 98	Mar 98 to Feb 99	Mar 99 to Nov 99	Dec 99 to Oct 00	Nov 00 to Mar 01
1	+0.7	+2.0	+10.1	+0.90	+12.6
2	+5.2	+3.3	+6.3	+2.2	+2.2
3	+2.3	+4.7	+1.6	+3.0	+5.8
4	+3.6	+4.1	+5.7	+0.8	+2.1
5	+4.0	+6.8	+2.6	+8.2	-4.3
<b>All</b>	<b>+3.9</b>	<b>+4.4</b>	<b>+4.7</b>	<b>+3.1</b>	<b>+1.6</b>

*These figures denote the estimated mean linear trend in each wave and period.*

#### 4.3.3 Conclusion

The evidence from this national study, examining the overall impact of the successive waves of NHS Direct, provides evidence that NHS Direct has been associated with a fall in the number of calls to general practitioner co-operatives. The size of the effect evident – a decrease of about eight calls per night for the average co-op – is considerably greater than the effect we found in our analysis of the first year of NHS Direct, perhaps because of the increase in call numbers over subsequent years.

By contrast, the current study suggests that the effect of NHS Direct on overall demand for ambulance services (a decrease of about two or three journeys per day for the average service) and A&E departments (an increase of one or two patients per day) continues to be negligible. We would note, however, that this does not necessarily mean there have been no effects on the operation of these services since it is likely that this “lack of effect” is actually the net result of two opposing effects: diversion of some patients away from these services, and of others towards them. Thus, there may have been an effect on the casemix of these services, which cannot be detected in the current study.

Some limitations of the current study must be acknowledged. In particular, data were obtained only from 35% of GP co-operatives thought to exist (though some may have merged or closed) and from 71% of A&E departments. Thus, there is potential for response biases which may have influenced the results in unknown ways. In addition, despite the very large number of service contacts used in the analysis, there is quite marked instability in the trends which are observed in each time period, particularly in the case of GP co-operatives and A&E departments. This instability inevitably introduces some uncertainty over the reliability of the conclusions we have drawn.

## **5. IMPACT OF NHS DIRECT: POPULATION SURVEYS**

### **5.1 Introduction**

Analysis of routine data from immediate care services can only give a partial picture of the impact of NHS Direct on patterns of health service use. We have sought to supplement this by undertaking an annual population survey in a number of areas served by NHS Direct, asking about where people have sought unplanned help or advice for health problems. This approach allows us to say something about whether the use of other services (for which large routine data sets may not be available) has changed since the start of NHS Direct, and also permits us to quantify the contribution of NHS Direct to all existing sources of help and advice.

### **5.2 Methods**

The core population survey covers the areas served by the original first wave NHS Direct sites. Random population samples in these areas – Milton Keynes, Chorley and Preston, and Newcastle, North Tyneside and Northumberland – were surveyed in 1998 (immediately before the introduction of NHS Direct) and in the following three years. The one exception is that in 2001 we were unable to include Milton Keynes as a survey area because the local health authority refused to provide a population sample. In addition, in the same year Sheffield was introduced as a new survey area although it is not included in the present analysis.

In each year, a random sample of 5,000 individuals (of all ages) was drawn from the health authority register in each area. In Newcastle and North Tyneside the health authority refused to provide a population sample from the start, so 3,000 names were drawn from the local electoral registers (which of course includes only those aged 18 and over) and added to the sample of 2,000 provided by Northumbria health authority.

Each person in the sample was sent a questionnaire and covering letter by the local health authority (or by the research team for the electoral roll sample), with up to two reminders to non-respondents at three weekly intervals. In 1998 and 1999 the first mailing was sent in February, with March being the month for the first mailing in 2000. In 2001, because of difficulties in obtaining samples the date of the first mailing differed between sites and ranged from March to June. The questionnaire did not ask for an individual's name or address but a unique identifying number was printed on each questionnaire so that non-respondents could be identified and reminders sent.

In each year the first three pages of the four page questionnaire remained identical. Respondents were asked whether they had sought help or advice for a health problem in the previous four weeks, however minor, and details of the contact made were sought. The last page of the questionnaire was constant for the first three years, and asked about access to a car and telephone, and any difficulties (emotional or physical) in using a telephone. In 2001, these questions were replaced with questions

on awareness of, access to and use of NHS Direct, NHS Direct Online and NHS Walk-in Centres.

### 5.3 Results

#### 5.3.1 Response rate and characteristics of respondents

Response rates were acceptably high in each year of the survey (Table 5-1).

Respondents were asked to record their age and sex. Over time, the age of respondents increased slightly, and in all years there was a higher proportion of female than male respondents.

**Table 5-1: Respondents to the population survey**

	1998	1999	2000	2001
Surveys mailed	15209	15147	15131	10018
Completed surveys returned	10620	10668	9958	6908
Response rate*	72%	73%	68%	70%
Male gender	47%	46%	46%	45%
Mean age	39.97	40.78	41.77	43.17

\*after removal of deaths and surveys returned by Royal Mail

#### 5.3.2 Incidence of health problems

Respondents were asked if they had suffered a health problem in the previous four weeks for which they sought help or advice. The proportion of people experiencing such a problem decreased very slightly over the four year period (Table 5-2).

Those who had sought help or advice in the previous four weeks were asked to list such contacts, and indicate which, if any, had been unplanned (in distinction to planned contacts, for example for a blood pressure check or a clinic visit).

“Unplanned” was defined as any contact which had not been arranged more than a day in advance. The number of such unplanned contacts fell slightly, in the same way that all contacts did.

**Table 5-2: Health problem in the previous four weeks for which help or advice was sought\***

	1998	1999	2000	2001	All years
Health problems	4029 (38%)	3770 (37.5%)	3546 (36%)	2458 (36%)	13803 (37%)
Unplanned health contacts	1776 (17%)	1688 (17%)	1486 (15%)	1029 (15%)	5979 (16%)

\* includes contacts with family/friends and health care professionals

#### 5.3.3 Use of healthcare services

Those experiencing such an unplanned episode were asked about all those from whom they sought help or advice. After adjusting for age, sex, area, and month in which the questionnaire was completed, statistically significant increases were found for two of the services across the four years: “helplines” and “someone else at my usual practice but not a doctor”. The increase in the use of “helplines” is probably a

simple indication of the use of NHS Direct. However, it seems unlikely that the increase in contacting someone other than a GP at GP surgery is attributable to NHS Direct, which does not refer callers to primary care professionals other than GPs. This change may be due to developments in the size and composition of the primary care team and the increase in services provided by GPs, rather than to the introduction of NHS Direct.

However, as the results in Table 5-3 show, there was no measurable change in the number of contacts made with a wide range of other services, including usual GP care, pharmacy, accident and emergency departments or dentists. (Use of out-of-hours general practice is probably best reflected in the row labelled “not usual GP”. Numbers here are too small to detect the change of around –6% per annum suggested by the wave analysis in the previous chapter.)

The analysis was repeated including only those respondents aged under five years, who account for about one in four calls to NHS Direct. In this restricted analysis, the use of helplines was the only statistically significant increase in service use.

Taken together, these findings add some weight to the assertion that NHS Direct has been associated with no measurable change in the overall volume of use of other NHS services.

**Table 5-3: Changes in service use for unplanned health events**

	Use in Year 0		Adjusted odds ratios*				p
	N	%	Year 0 1998	Year 1 1999	Year 2 2000	Year 3 2001	
Usual GP	1155	10.9	1	0.939	0.983	1.004	0.539
Family	880	8.3	1	1.024	1.117	1.047	0.464
Chemist	853	8.0	1	0.971	1.058	1.084	0.583
A&E/999/Ambulance	148	1.4	1	1.285	1.132	1.034	0.136
Someone else at GP practice but not a doctor‡	127	1.2	1	1.044	1.603	1.801	<b>0.028</b>
Not usual GP	114	1.1	1	1.093	0.942	1.333	0.324
Outpatient clinic	104	1.0	1	1.134	1.162	0.780	0.097
Dentist	92	0.9	1	0.904	0.918	0.768	0.744
Helpline†	72	0.7	1	1.751	1.599	1.968	<b>0.001</b>
Hospital Admission	58	0.5	1	1.268	0.938	0.789	0.265
Complementary Therapist	48	0.5	1	0.881	1.555	1.272	0.212
Physiotherapist	44	0.4	1	0.994	1.241	1.133	0.873
Social worker	16	0.2	1	1.230	1.069	0.382	0.256
Other Mental Health Services	15	0.1	1	0.884	0.761	0.975	0.924
Community Psychiatric Nurse	14	0.1	1	0.814	0.479	1.288	0.377
Community Nurses	14	0.1	1	1.933	1.749	1.193	0.200
Minor Injuries Unit	4	0.0	1	1.560	2.522	0.994	0.499

\*adjusted for age, sex, area, and month when the questionnaire was completed

† 95% confidence intervals for helpline coefficients: 1999 (1.3, 2.4), 2000 (1.1, 2.4), 2001 (1.3, 3.1)

‡ Over half of such contacts relate to a community nurse/practice nurse. 95% CIs: 1999 (0.8, 1.3), 2000 (1.1, 2.3), 2001 (1.2, 2.7)

### 5.3.3.1 Number of contacts made for an unplanned health event

There has been intermittent concern since NHS Direct was established that it would lead to an increase in the number of contacts people make for any given health episode in which they seek unplanned help or advice. For example, it might add to the length of the “pathway to care” through referring patients inappropriately, or act cautiously by suggesting multiple contacts.

However, this does not seem to have occurred in any measurable way in the populations surveyed. In each for the four years, the mean number of contacts with healthcare providers for the most recent unplanned health event recalled was 1.4 contacts per episode. This did not change over the period examined.

### 5.3.3.2 Use of NHS Direct during unplanned health events

Respondents were not asked explicitly whether they had made use of NHS Direct as a result of an unplanned health event, but they were asked for the details of all healthcare providers which they had contacted during the episode. Table 5-4 shows the number of respondents who clearly indicated that they had used NHS Direct (giving the name or telephone number of the service), and those who may have used it (for example, describing an “NHS helpline”).

This data suggests that, over the first three years of NHS Direct in the first wave sites, there has been some small increase in the use of NHS Direct in unplanned episodes of care. However, it is also clear that during this period NHS Direct accounted only for a small proportion of all such episodes.

**Table 5-4: Use of NHS Direct for unplanned health events**

	1999	2000	2001
Unplanned health events	1688	1486	1029
Definite contacts with NHS Direct	39	43	43
Possible contacts with NHS Direct	85	81	60
Proportion of events for which NHS Direct was used (range)	2% to 5%	3% to 5%	4% to 6%

## 5.4 Conclusions

In this analysis we have brought together data from population surveys carried out in the same way in four consecutive years, spanning the introduction of NHS Direct, in order to examine the unplanned use of a wide range of immediate care services.

The findings show no measurable impact of the new service on population rates of use of existing services, with the possible exception of an increase in the use of services other than GPs at local surgeries (though we regard this as an unconnected phenomenon). In addition, there has been no measurable impact on the mean number

of contacts with health care providers which people make during unplanned episodes of care. These results are not surprising in the light of the finding that people use NHS Direct in only about one in twenty unplanned episodes of care, and it is clearly possible that, were rates of use of NHS Direct to increase, measurable effects on other providers might become apparent.

## 6. “APPROPRIATENESS” OF TRIAGE OUTCOMES

### 6.1 Introduction

The idea of “appropriateness” is of central importance to the purpose of NHS Direct. The overall aim of the service, both in original conception and in the way it has developed since its inception in 1998, is to benefit both the users and providers of health care by guiding callers to the service most *appropriate* to their particular health problem. Indeed, this is explicit in the published aims of NHS Direct (see section 2.1, Policy background).

It must therefore be a central issue of an evaluation, at some point, to attempt to assess the degree to which NHS Direct is in fact able to direct callers to “appropriate” services, especially since there has been considerable public and professional debate both about the possibility of doing this reliably through telephone triage, in theory, and whether it actually occurs in practice.

### 6.2 Conceptual issues

#### 6.2.1 The idea of appropriateness

Attempts to assess appropriateness in health care have a long and, some would argue, unconvincing history. The idea of appropriateness embodies notions not only of what treatment, advice or referral is “correct”, but what is right under the circumstances, “all things considered”. It clearly includes the idea of effectiveness, but for some would also include the idea of cost-effectiveness and even fairness. In some instances it seems to merge into the idea of “avoidability” (for example, of referrals, attendances, or admissions).<sup>15</sup> It is also clear that what is judged to be appropriate may depend upon the perspective of the person making the assessment. The idea itself is slippery, quite apart from any attempt to turn it into a measurable quantity.

A number of attempts to produce a clear definition of appropriateness have been made. For example, a well known definition from the US Rand Corporation states that care is appropriate when:

*for an average group of patients presenting to an average US physician... the expected health benefit exceeds the expected negative consequences by a sufficiently wide margin that the procedure is worth doing... excluding considerations of monetary cost.*<sup>16</sup>

A UK working group which considered this definition felt that it lacked two further dimensions – the individuality of the patient, and the availability of health care resources. Their suggested revised definition was so long and involved so many different dimensions that, for evaluative purposes, it is effectively useless.<sup>17</sup>

In practice, the attempt to apply the Rand Corporation definition, using expert panels of assessors, has raised many questions about how such judgements should be

made and what they can be taken to mean.<sup>18</sup> Assessments of appropriateness may depend upon the supposed aims of care, the risks and benefits which are to be considered, beliefs about the effectiveness of interventions, the weight given to different possible outcomes, judgements about the value of an outcome in relation to the effort or cost involved, the alternative courses of action available, whose point of view is to be represented, the weight given to the wishes of the individual patient, and indeed the composition of the panel making the assessments.<sup>19</sup>

In view of this, further conceptual clarification may be helpful. One recent review of the topic, in the context of prescribing practice, draws attention to the need to distinguish between the *process* of decision-making and its *outcome*.<sup>20</sup> Thus, *rational* prescribing (a decision arrived at in a logical way, processing all available relevant information) is distinguished from *appropriate* prescribing (in which the outcome of the decision, however reached, is judged in some sense as “right”). Taking this idea further, one can attempt to distinguish a number of different approaches which might be taken in operationalising the idea of appropriateness for evaluative purposes. For example:

- Decisions made by following some agreed or formal process
- Decisions regarded by others (“experts”, patients, or other service providers) as “good practice”
- Decisions which are justified, in some way, by subsequent events or outcomes

These approaches will be examined in the context of NHS Direct, and in the light of research to date.

#### 6.2.2 Approaches taken in evaluations of NHS Direct

A number of evaluations of NHS Direct have been undertaken in which the rightness or appropriateness of the triage decision or self-care advice given has been assessed.

To begin with, it might be argued that NHS Direct algorithms already embody agreed good practice and so the triage decisions and self-care advice which follow from their use must, of necessity, be appropriate. If this were the case, then evaluation would simply be a case of monitoring the use of the algorithms to ensure that in practice they were being used as intended.

Unfortunately, things are not quite so simple. First, in the absence of some kind of evaluation we do not know for sure that the algorithms do embody good practice. To date, NHS Direct has used four different sets of algorithms, all regarded as good practice, but which lead to quite different outcome distributions.<sup>14</sup> It is unlikely that all are equally “appropriate”. Second, the choice of which algorithm to use in a call is an interpretive act, a matter of the nurse’s judgement.<sup>21</sup> Third, algorithms may not be available for all problems which present to the service.<sup>5</sup> Fourth, since the nurse can

over-ride the computer endpoint, she must interpret its appropriateness and applicability for herself, in deciding whether and how to present it to the caller. Fifth, there may have to be negotiation between the nurse and caller over what will be done next. Sixth, the caller may or may not follow the course of action recommended (although a large majority say that they do).<sup>6</sup>

Thus it is apparent that, in reality, a pragmatic evaluation of the appropriateness of the outcomes of the triage process is required.

Two published evaluations have assessed NHS Direct advice by reference to the views of other health care professionals (“experts”). In the first, a small study conducted by the Consumer’s Association, three actors posing as callers presented three problems to the service.<sup>10</sup> The calls were recorded and assessed by a panel of “experts”, consisting of three GPs, a nursing professor and a pharmacist. No explicit criteria for assessment were given, and the study was criticised for weak methodology.<sup>11</sup>

A second study taking a similar, though more rigorous, approach was conducted in the context of the national evaluation of walk-in centres.<sup>22</sup> In this study five clinical scenarios were presented to services by actors trained in their portrayal. The scenarios were each presented on 99 occasions to walk-in centres, NHS Direct and general practices. Clinical care was assessed by reference to prospectively determined standards, derived from a Delphi process, and scored. Comparisons were then made between the three types of service. The results of this study suggested that, taking the five scenarios as a whole, NHS Direct compared well with general practice on history taking but less well on diagnosis, advice and treatment.

This kind of detailed approach to the content of the clinical consultation has much to offer in understanding, for specific conditions, how consultations go well or badly. However, it has limitations in providing an assessment of NHS Direct as a whole since the approach is inevitably scenario-specific, and the performance of NHS Direct (as of other services) may differ markedly from one scenario to another.<sup>23</sup>

A small number of studies have also been carried out in which callers to NHS Direct who are referred to immediate care services have then been “second triaged” by the subsequent provider. By comparing the initial triage with their own, providers have sought to determine whether the first triage was, or was not, appropriate. For example, in the only study of this type which has been published in a peer reviewed journal, 91 attenders at an accident and emergency department who had been routed to the 999 service by NHS Direct were compared, in terms of their A&E triage category, with 232 people who had “self-referred” to 999.<sup>24</sup> The study found that there was a similar distribution of triage categories between the two groups, from which the conclusion was drawn that “severity of illness is equally well assessed by self as by NHS Direct”.<sup>24</sup>

There are at least three difficulties with this “second triage” approach to assessing appropriateness. First, there is no certainty that the second triage is, in some sense, “better” than the first. Just as we have no clear idea of what appropriateness means, we also have no “gold standard” for its measurement. We simply have two assessments, which may agree or disagree.

This leads to a second problem, which is that of regression to the mean.<sup>25</sup> Suppose that there is some measurable attribute of case severity or urgency, and that services are able to measure it, albeit with some risk of error. Measured on the first occasion (by NHS Direct), some patients will be allocated to a particular category of severity. A second measurement of the same attribute will then be carried out by the service to which they are sent. Inevitably, because of measurement errors (either on the first or second occasion), some of these will be allocated lower or higher categories of severity on the basis of the second measurement. *On the average*, the second triage of urgent cases will then be less urgent than the first triage: it will have “regressed to the mean”, but this cannot be taken to indicate that cases have been consistently over-triaged.

Third, there is the possibility that the service to which NHS Direct refers is motivated by a desire to show that it is being asked to deal with “inappropriate” cases. Thus, in the absence of effective blinding, there may be a measurement bias towards showing that NHS Direct-referred cases are inappropriate in some way. This seems to have occurred in at least one publicised case in which a joint audit of referrals was being conducted by NHS Direct and a local GP co-operative.<sup>26</sup>

A different approach to assessing appropriateness is to seek the views of callers themselves. While we are not aware that callers have been asked to make this assessment to date, a number of surveys have asked callers to rate their satisfaction with the service, or the helpfulness of the advice offered. In general such ratings have shown a high level of satisfaction (consistently above 90% “satisfied” or “very satisfied”)<sup>4 6</sup> suggesting that, in the caller’s view, advice given in the large majority of calls is likely to be “appropriate”, although of course the services to which they are referred may not agree.

### 6.2.3 The idea of necessariness

The approach taken in the present study attempts to avoid many of the difficulties in assessing appropriateness described above by recasting the question into one of “necessariness”. Rather than ask whether the advice given was appropriate, we ask: *did the first contact made by the patient following NHS Direct advice prove to be necessary, more than necessary or less than necessary?* In taking this approach we restrict ourselves to the triage component of NHS Direct’s advice. We do not attempt here to assess the rightness or quality of the self-care advice given.

The essence of this method is to focus on what happened – on the processes of care experienced by the patient. We then judge, from the care that the patient received, whether the patient's contact with that service was indeed necessary, or whether they could have been cared for elsewhere. Such an approach has previously been used in assessing whether attenders at an accident and emergency department needed the facilities of the department or could in fact have been cared for by their general practitioner.<sup>27 28</sup>

This method has the virtue that, in principle, it can be applied in an explicit (transparent), consistent and reproducible way. It also has the advantage that large numbers of cases can be classified much more rapidly than using the "expert" approaches described above. In addition, in principle the rules of classification may be varied and cases re-analysed relatively easily, to allow sensitivity analyses.

However, the approach relies upon a number of central assumptions which, though reasonable, are not self-evident. These are as follows:

1. That a "ranking of care" can be defined – in other words, that services can be ordered in some way that makes sense: perhaps in terms of the skills and facilities available, the cost of a contact, or the scarcity of contact time available from the service. Though one might argue about the exact basis for the ranking and about the precise order in which services should be placed (as we did during this study) the idea that services are ordered in some sort of way is a common one in the minds of professionals and the public, and it is therefore not unreasonable to try to make such an order explicit.

For the purposes of this study, we defined a ranking of care (from top to bottom) as accident and emergency department, general practice, nurse-led services (such as minor injury units and walk-in centres), community pharmacy services, and finally self-care. We were unable to agree on how dentists might be placed in this category, so cases contacting a dentist first were analysed separately.

Although patients do contact dentists and a range of other services as a first port of call, the number who do so is tiny in proportion to those contacting the services listed in the ranking.

2. That, other things being equal, patients should be directed to the "lowest" level of care able to meet their needs.
3. That services are assumed to do all that is necessary (including referral, if need be), and nothing which is not necessary, to meet the patient's needs. Clearly this may not be the case on every occasion, but we assume it is the case often enough to be true as a generalisation. From this rule it follows that we take the processes of care which occur as the processes of care which were needed, and we do not try to "second guess" the clinician who saw the patient.

4. That a patient will seek further help if their problem remains unresolved or worsens, and they are able to do so. Thus, we assume that in general a person will not remain quietly at home with a broken limb, worsening breathlessness or increasing abdominal pain, even if advised to do so, but will seek further help from some other agency. The corollary is that if the patient does not seek any further help, we assume that their problem has either resolved or has lost its urgent character.

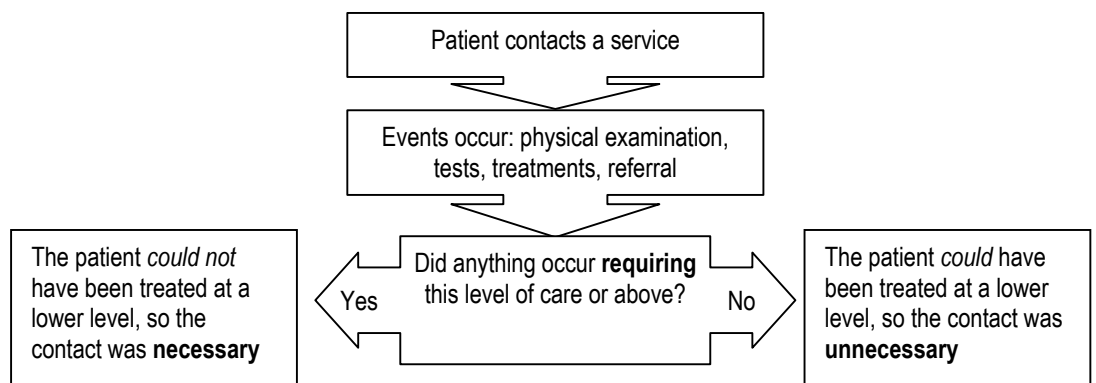
Of course, this assumption does not hold if the patient is unable to seek further care because they died before they could do so. The possibility of this eventuality must be checked for in some other way.

5. That all relevant processes of care are recognisable to the patient and can be reported for analysis. While this is likely to be true for most physical examinations, investigations, treatments and referrals, it is of course the case that patients are often examined through observation (for example, of facial appearance, respiratory rate, gait or mental state) without being aware of the fact. In cases where such examination is the only determining event which occurs, a process-based approach will misclassify the contact.

Given these assumptions, the necessariness of a patient's contact with a service can be assessed by examining what happened during that contact – the examinations, tests, treatments, recommendations and referrals made. If anything occurred which could only be provided by that level of care, or a higher level, we rate the contact as *necessary* (though not necessarily *sufficient*). Conversely, if nothing happened which required that level of care, then we rate the contact as *unnecessary*, since the patient could have been cared for at a lower level of care.

This is shown diagrammatically in the figure below.

**Figure 6-1: Assessing the "necessariness" of a contact**

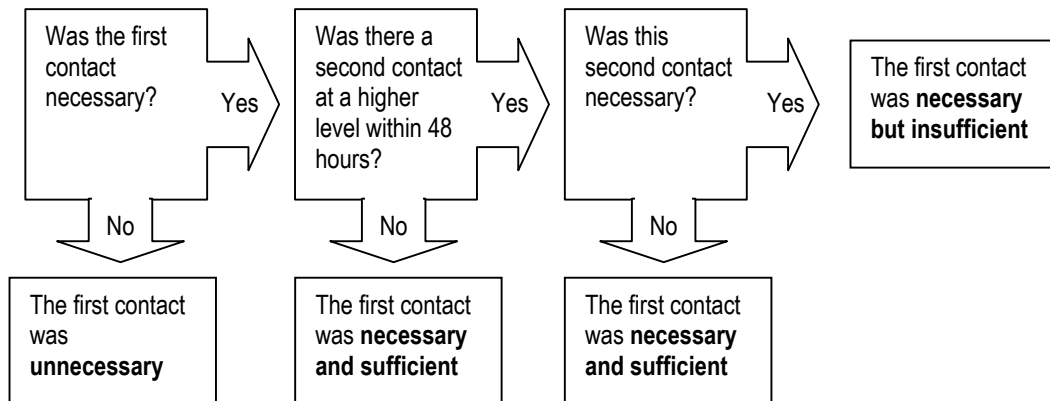


If it is the case that the patient's first service contact is necessary (i.e. justified by the events which occur), then we ask whether a second contact with a "higher level" service followed the first, within some time period such as 48 hours. If it did not, then it

follows from assumption 4 above that the first contact was both *necessary* and *sufficient*. On the other hand, if a second contact does follow then we ask, using the same method as before, whether this contact was necessary. If it was, then it follows that the first, while necessary, was *insufficient*. If it was not, then it does nothing to undermine the sufficiency of the first contact.

This further complexity is shown diagrammatically below.

**Figure 6-2: Assessing the "sufficiency" of a contact**



Thus, at least in principle, by examining the processes of care which occur in up to two contacts following the call to NHS Direct, we can classify the first contact made by each patient in one of three ways: necessary and sufficient, unnecessary or necessary but insufficient.

### **6.3 Methods**

In summary, the study involved defining a ranking of care (given above) and listing observable processes of care (events) which a patient might report as having occurred at a service. We developed a set of "rules" determining which processes of care were available at which service level. We asked a large number of callers to NHS Direct to report their processes of care, by means of a postal survey. Our rules were then applied, either by computer or by clinician, to the processes of care reported. For each caller, the purpose was to classify the necessariness of their first contact after their call.

Subsequently, by examining whether the action the caller took was also the action which NHS Direct had advised (i.e restricting the analysis to compliers), one can draw conclusions about the degree to which NHS Direct advises contacts which prove to be necessary ones.

### 6.3.1 Definition of events and rule set

#### 6.3.1.1 *Initial clinical event rules*

An initial set of possible events was drawn up and developed further through discussion with the clinical panel (see below for details of the panel). These events consisted of all those processes of care which we felt would be reasonably common, recognisable by the patient as having occurred or not, and for which there was space on the postal survey, which we wanted to limit to a maximum of eight pages. The initial list was used in the pilot survey, and further events were added in the light of this. The final list of events included in the survey is shown in Table 6-1, below.

The list included a total of 39 events, of which 32 were provided on the survey as simple “tick box” (binary) events. The remaining 7 were provided as boxes inviting a free text description of the event.

Each of the 32 binary events was allocated to a level, which was the lowest level at which the panel felt one could be reasonably sure that the event would be available, assuming that all services were available in all localities at all times (although we vary this assumption later in the analysis). For example, a dressing, sling or support bandage could be reliably supplied at the level of nurse-led services (MIUs and walk-in centres) and above. Although one might be lucky enough to be obtain a support bandage at a pharmacy, this could not be relied upon. These levels are also shown in Table 6-1.

The seven events for which a free text description was invited clearly could not be allocated a level in advance, although some could be coded to simple binary events, as explained below. Where necessary for determining the classification, the remaining free text data was presented to the panel (see below), as indicated by “review of free text” in the table.

**Table 6-1: Initial events included in the postal survey**

Event	Lowest level of care at which this event is reliably available
Physical examination	
No physical exam	Self care
Not sure if physical exam	Self care
Chest examined	GP or emergency doctor
Abdomen examined	GP or emergency doctor
Skin examined	MIU/WIC
Eyes examined	GP or emergency doctor
Ears, nose or mouth examined	MIU/WIC
Rectal or vaginal examination	GP or emergency doctor
Limbs examined	MIU/WIC
Other examination was done	Review of free text
Tests	
No tests	Self care
Not sure if any tests were done	Self care

Blood tests were done	GP or emergency doctor
An ECG was done	GP or emergency doctor
X-rays/scans were done	A&E
Urine tests were done	MIU/WIC
Other tests were done	Review of free text
Treatments	
No treatment	Self care
Not sure if treatment	Self care
Drip or injection there and then	Review of free text
Medicine or prescription to take later	Review of free text
A dressing, sling or support bandage was provided	MIU/WIC
Plaster was applied	A&E
Stitches were used to close a wound	MIU/WIC
Steristrips were used to close a wound	MIU/WIC
Glue was used to close a wound	MIU/WIC
Other treatment was given	Review of free text
Advice	
Advice/information was given	Review of free text
Follow up	
Follow up appointment was made	Justifies contact at that level
No further care was advised/arranged	Self care
Not sure about further care	Self care
Hospital admission was arranged	A&E (but see additional rules below)
A&E was advised	GP or emergency doctor
MIU/WIC was advised	Pharmacist
Hospital clinic was arranged	GP or emergency doctor
Advised to see GP	Pharmacist
Advised to see pharmacist	Self care
Advised to see a dentist	Pharmacist
Other care was advised/arranged	Review of free text

### 6.3.1.2 Derived events

Among the additional processes of care mentioned in the free text boxes provided, some were listed very commonly, such as the examination of the neck or shoulder, the taking of temperature or pulse, or the administration of oxygen. A further 26 of the most common events mentioned were added to the list of binary events and assigned a level by the panel in the same way as the initial events, as shown in Table 6-2. During data cleaning, free text responses were read by a qualified medical practitioner (JM) and recoded to one of these events where appropriate.

**Table 6-2: Additional events derived from free text data**

Event	Lowest level of care at which this event is reliably available
Physical examination	
Back	GP or emergency doctor
Ribs	GP or emergency doctor
Pulse/heart rate	MIU/WIC

Event	Lowest level of care at which this event is reliably available
Breast	GP or emergency doctor
Conscious level	MIU/WIC
Fontanelle	GP or emergency doctor
Head injury	MIU/WIC
Face	GP or emergency doctor
Gums	MIU/WIC
Neck	GP or emergency doctor
Shoulder	GP or emergency doctor
Reflexes	GP or emergency doctor
Tests	
Temperature measured	MIU/WIC
Blood pressure	MIU/WIC
Blood oxygen test	A&E
Blood sugar test (BM sticks)	MIU/WIC
Stool test	MIU/WIC
Heart rate monitor	A&E
CT scan	A&E
Treatments	
Gas and air	MIU/WIC
Nebuliser	A&E
Oxygen	A&E
Nose cauterised	GP or emergency doctor
Nail trephine	MIU/WIC
Eye wash for trauma	A&E

### 6.3.1.3 Drug-related events

Two questions in the survey invited the respondent to report on drug treatments they had received, whether as a “drip or injection there and then”, or as a “medicine or prescription to take later”. These questions generated a wide range of free text data.

This free text data was recoded by a qualified medical practitioner (JM). Responses to the former question (drip or injection) were recoded as either intravenous infusion or bolus intravenous/intramuscular injection events, and allocated a service level as shown in Table 6-3. Responses to the latter question (medicine or prescription) were recoded as recognisable drug names.

A list of all licensed medicines used in the UK was obtained from the Royal Pharmaceutical Society of Great Britain, which included 6343 separate drugs and preparations.<sup>29</sup> For each drug, the list gave its legal prescription status, which determines how the drug may be supplied, as GSL (general sales list – i.e. an over

the counter medication), P (pharmacist only), and POM (a prescription only medicine). In many places, however, nurse-led services have agreed limited drug formularies of medications which may be supplied in a limited quantity by nurse practitioners. To account for this, we obtained such formulary lists from the Sheffield minor injuries unit and the Sheffield walk-in centre. Where a drug appeared on these lists we altered its status to N (nurse-prescribable medication). In some cases this status applied only to a limited group of patients, such as those aged 12 and over. This was also coded into the drug list.

Each drug listed by a patient in describing their care was then matched, where possible, to its entry on the drug list, and new events were created (as shown in Table 6-3) to reflect the status of the medication which had been used.

**Table 6-3: Drug events derived from free text data**

Event	Lowest level of care at which this event is reliably available
GSL medication	Self care
pharmacy medication	Pharmacist
nurse-prescribable medication	MIU/WIC
doctor-prescribable medication	GP or emergency doctor
IV infusion	A&E
stat IV or IM drug	GP or emergency doctor
A&E prescribable treatment	A&E

#### 6.3.1.4 Additional rules

In the course of early panel attempts to apply event rules to cases, a small number of additional rules and clarifications were agreed, as follows:

- a) **Two contacts at the same level:** where the patient reports two contacts at the same level (for example, seeing an emergency GP at night followed by their own GP the next morning), then if the second contact is justified by events, so is the first.
- b) **Acceptable routes of admission to hospital:** if a patient is admitted directly to hospital by a GP or via A&E, or a GP refers the patient to A&E which then admits the patient, this justifies the contact in each case.
- c) **Children under one:** The panel felt that “ill children” under one generally required assessment at GP level or above. Thus, a GP contact by an ill child under one would be justified, even in the absence of other events. This rule was the only departure from our process-based approach.
- d) **Stopping a prescribed medication:** If a prescribed medication was stopped, this would justify a GP contact.

### 6.3.2 Survey of recent patients

Data on the processes of care experienced by patients following calls to NHS Direct were collected by means of a postal survey.

#### 6.3.2.1 *Survey development and piloting*

An initial survey was developed and piloted with a random sample of 500 patients of NHS Direct, in a single NHS Direct site. The survey was mailed by NHS Direct within 10 days of the call, and one reminder only was sent to non-respondents. Each survey was sent with a covering letter from NHS Direct, and a reply-paid envelope. In all, 254/492 surveys were returned (8 were returned by the Royal Mail as uncontactable), a response rate of 52%.

The main purpose of the pilot was to test the layout and ease of completion of the questionnaire. Examination of the responses showed that the questionnaire was being completed as intended, and suggested the need for additional process of care events to be added.

The pilot survey was also used to test two methods of seeking patient consent to examine medical records relating to the call. In method A, we added an extra question at the end of the survey which asked the respondent to give their consent to the research team to access their medical records relating to the call. In method B, this question was omitted (since we feared it may lower the response rate) and instead a letter was sent to respondents as soon as their completed survey was received, thanking them for returning it. The letter also asked for consent to be given, by completing and returning a tear-off slip which contained the same question as that included on the survey for method A. A pre-paid envelope was included.

In practice, the response rates to both surveys were identical. Using method A, we gained consent from 70% (85/127) respondents to the survey. Using method B, we gained consent from only 27% (29/106) of respondents whom we contacted by follow-up letter. We therefore decided to mail the main survey with the consent question included.

#### 6.3.2.2 *Survey sample*

Three NHS Direct sites were selected for the study, and all agreed to participate. One site was located in the south-east, one in the west and one in the north-west of England, and all had been using the NHS clinical assessment system for some months. Table 6-4 indicates the availability of selected immediate care services in each of the NHS Direct areas. It may be noted that nurse-led services (minor injury units and walk-in centres) are not equally available across these sites, a point to which we return in the analysis below. We assume that GP and community pharmacy services are readily available to patients in all areas.

Each site was asked to identify a random sample of 2,000 patients from NHS Direct call logs who had used the service during a single week in January/February 2002. Each patient was allocated a study number by the site. Patient identities were not released by the sites to the research team, to maintain confidentiality.

**Table 6-4: Selected immediate care services available in study areas**

Site	Location	Selected immediate care services within area served		
		A&E	Minor Injury Units	Walk-in centres
1	Southern England	11	4	0
2	North-west England	19	8	12
3	West England	10	23	4

We chose to survey patients rather than callers (these are different people in about half of all calls) since contact details are reliably recorded for the former but not the latter, and the caller may not know the subsequent processes of care which the patient undergoes. We also decided not to restrict calls to those for which NHS Direct records indicated that the call had undergone nurse triage, since this may not be reliably recorded, and there may not be a hard and fast distinction between “information” calls and “advice” calls in any case (as the *Health Which?* assessment highlighted<sup>10</sup>).

As in the pilot, surveys were mailed by each site as soon as practicable after the call, along with a covering letter from NHS Direct and a reply-paid envelope. Responses were returned directly to the research team. Up to two reminders were sent to non-respondents at three-weekly intervals. The survey of patients was carried out during February and March 2002. Unfortunately, in one site the data extract of patients proved to be faulty so that the survey had to be repeated, based on calls made between 10 March and 24 March 2002.

The eight page survey asked respondents to consider the most recent call they had made (or had been made on their behalf) to NHS Direct in the previous four weeks. Questions were included about the nature of the problem for which advice was sought, the triage and self-care advice offered by NHS Direct, whether the advice was followed and how easy this was to do (see Appendix 1).

The bulk of the survey asked respondents to recall the contacts with services they had made following the call and, for the first two contacts, to indicate the processes of care which had occurred, as discussed above. Finally, respondents were asked whether they considered the advice had been appropriate and helpful, and whether NHS Direct had assessed the urgency of their problem correctly.

Data from all completed surveys were entered into a Microsoft Access database for cleaning and analysis.

### 6.3.2.3 *Non-responders*

We noted above the one reason for non-response may be serious illness or death following the call to NHS Direct. In order to examine this possibility we asked each NHS Direct site to write to the GPs of a random sample of 100 non-responders some months after the survey. GPs were asked to indicate on a tear-off slip whether or not their patient had been seriously ill or had died during the two week period following their call to NHS Direct. The slip was to be returned directly to the researchers, and in order to maintain confidentiality included the patient study number but no other identifying details.

### 6.3.3 Classification of cases

Respondents who denied having called NHS Direct in the previous four weeks were excluded from further analysis. Of those who said they had called, many had not made any further contact with health services following the call. Following assumption 4, above, we classified such lack of contact as *sufficient*.

Those who had made further contacts were classified in up to three stages, as follows:

**Stage 1:** The rule set agreed by the panel was applied to each contact of each case in turn, by the computer. Where contacts could be justified by one or more binary events (for which unambiguous rules exist), they were marked as justified. Where no such justifying binary event was reported and no free text data was reported either, the contact was marked as unjustified. Where no justifying binary event was reported but some free text data was present, the case was marked for review by the panel. The computer considered all the initial, derived and drug-related events described above. Using this system the computer rated all cases as either *necessary* (justified) *and sufficient*, *unnecessary*, *insufficient*, or for review by the panel.

**Stage 2:** Cases marked for review were presented to panel members individually, such that each case was seen by two panel members. Each panel member was asked to read the free text data and consider the level at which the processes of care reported could be provided – in other words, to apply the same logic of classification as that applied by the computer. Panel members were explicitly asked *not* to make judgements about the care they felt would have been “appropriate” for the respondent’s problem, but only to judge, on the basis of the events reported, whether the contacts made by the respondent had proved to be necessary or not.

Panel members were aware of the problem presented by the caller, but all panel members remained unaware both of the advice that the respondent reported having been given by NHS Direct, and of whether they had followed it.

Where the classification of both panel members agreed, this was taken as the final rating. For cases where they disagreed, they were asked to discuss the case and

seek agreement if possible. If they could not easily agree, or they felt the case raised additional issues for the methodology (such as a new rule, or clarification of an existing rule) the case was referred to the full panel for discussion.

**Stage 3:** A small number of cases were referred to the full panel for discussion, clarification of rules or suggestion of new rules. The panel sought consensus on the classification of the case in question. In case of further disagreement, a vote of 5-1 in favour of a particular classification was taken as final.

#### *6.3.3.1 Expert panel*

At the start of the study an expert panel of six clinicians was formed to aid in the development of the methodology and the classification of contacts. The composition of the panel was as follows:

- Two emergency medicine consultants with extensive clinical research experience;
- Two general practitioners: one also a medical director of NHS Direct, the other with previous management experience in a GP co-operative, a walk-in centre and an NHS Direct site;
- One NHS Direct nurse consultant with extensive clinical research experience;
- One walk-in centre and NHS Direct nurse manager.

The panel was asked to perform various roles during the conduct of the study. It reviewed the initial set of events and associated levels of care, and made suggestions for improvement. Panel members assessed cases marked by the computer as “for review” as set out above, and the panel clarified and occasionally created new rules as the study progressed, which were then applied to all cases.

It is important to note that the panel was specifically guided away from making judgements about appropriateness, but instead was asked to restrict its attention to the consistent application of explicit criteria to reported facts about processes of case. Thus, the aim was to avoid the subjectivity and lack of transparency associated with many previous attempts to judge appropriateness. In addition, the panel remained unaware at all times of the actual advice given by NHS Direct.

It may be noted that some members of the panel were associated with NHS Direct, raising the possibility of bias in their assessments. It would have been possible to compose a panel entirely of clinicians without such associations. However, on balance we felt that the advantages of including these clinicians on the panel outweighed the disadvantages, since they were familiar with all the immediate care services and systems involved, facilitated access to NHS Direct sites and staff, and would bring credibility to the results from the point of view of NHS Direct itself. Assessment bias would be unlikely since the clinicians were working in pairs in which each NHS Direct

clinician was paired with a non-NHS Direct clinician, cases were being rated using explicit rules, and, crucially, all raters were blind to the NHS Direct advice given or whether the patient had complied. Thus, they would not be able to predict how their rating of any individual call would affect the overall assessment of NHS Direct.

#### 6.3.3.2 *Service scenarios*

One of the assumptions of the “ranking of care” set out above is that all services are universally available in all places. In the case of nurse-led services, in particular, this is clearly not the case.

In order to take this into consideration, cases from one site were classified three times, making different assumptions about available services each time. Under scenario A, cases were classified assuming all services were available. Under scenario B, it was assumed that although nurse-led services were available, no nurse prescribing was permitted. Under scenario C, there were no nurse-led services at all. For each scenario, the rules were changed slightly so that the “lowest available level” for each event considered only available services. In other respects, cases were classified in the same way as previously.

#### 6.3.3.3 *Compliance with NHS Direct advice*

Using the approach set out above, we are able to classify cases according to the contacts which the patient actually makes, on the basis of the processes of care which occur. In a pragmatic sense, then, we may say we are evaluating the effects of NHS Direct triage, whether or not people follow the advice they are given.

In a more explanatory analysis, however, we would wish to know whether the triage advice given is “necessary” – i.e. refers the patient to a necessary service. In this cases we must restrict the analysis only to calls in which the patient complied with the advice. Clearly, it is only in these calls that we can use information on the patient’s actions to draw conclusions about the advice they were given. In general, if the patient has not followed the advice, then the events which follow their actions cannot be used to assess the advice offered.

In the current study, the compliance of the patient was determined in one of two ways: first, (method A) simply by the answer to the question “Did you act on the advice?”; and second, (method B) by comparing the advice the patient says they were given with the actions they say they took, and creating a derived compliance variable. Using the “ranking of care” set out above we coded each patient’s actions as being as advised, “more” than advised (i.e. their first contact was higher in the ranking) or “less” than advised (i.e. their first contact was lower in the ranking).

(A third possibility is to compare reported actions with the CAS log, but this was not available to us; a fourth is to examine the records of local health services to see whether the patient attended, but this is time-consuming and brings other difficulties.)

## 6.4 Results

### 6.4.1 Survey of recent patients

In all, 5921 patients were surveyed across the three sites. Of these, surveys to 46 patients were returned by the Royal Mail as unknown at the address, and 11 were returned because the patient had died. Excluding these, 3468 responses were received, a response rate of 59% (3468/5864). However, 89 responses were unusable because the respondent had completed few or none of the questions. Thus, 3379 responses were available for analysis.

#### 6.4.1.1 Characteristics of respondents and calls

The age and gender distribution of survey respondents is shown in Table 6-5, below. In all, 35% (1195/3379) respondents were male and 64% (2173/3379) female, and almost all respondents were registered with a GP (99%, 3334/3379).

**Table 6-5: Age and gender of survey respondents**

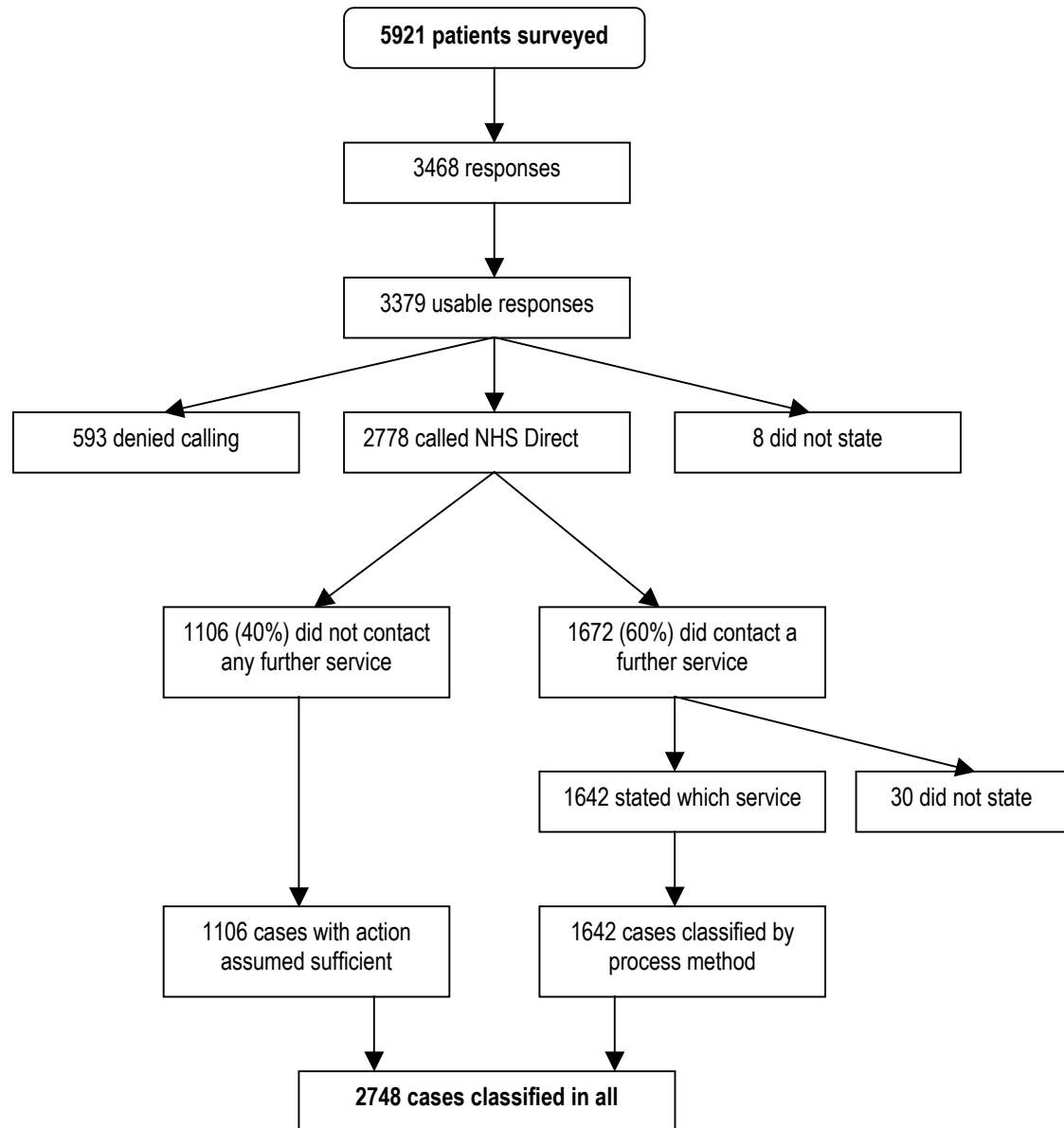
Age group	Male		Female	
	n	%	n	%
0: 9	517	43%	549	25%
10: 19	90	8%	163	8%
20: 29	81	7%	359	17%
30: 39	112	9%	391	18%
40: 49	94	8%	216	10%
50: 59	90	8%	188	9%
60: 69	95	8%	112	5%
70: 79	63	5%	96	4%
80 plus	33	3%	69	3%
No age stated	20	2%	30	1%
<b>All respondents</b>	<b>1195</b>	<b>100%</b>	<b>2173</b>	<b>100%</b>

*Note: 11 respondents who did not state their gender have been omitted from this table*

Asked about their contact with NHS Direct in the previous four weeks, 76% (2574/3379) said they had called NHS Direct themselves, or a call had been made on their behalf, and a further 6% (204/3379) had tried to contact their own doctor but had been put in contact with NHS Direct, which may have been through call switching or by means of a message on the GP answerphone. A substantial number (17%, 593/3379) denied having called NHS Direct. This may have been because they could not recall the contact, someone else had phoned without their knowledge, they had called for GP out-of-hours advice but had been connected to NHS Direct without realising, their details had been incorrectly recorded, or because delays in sending, receiving or replying to the survey may have meant their last call was longer ago than four weeks.

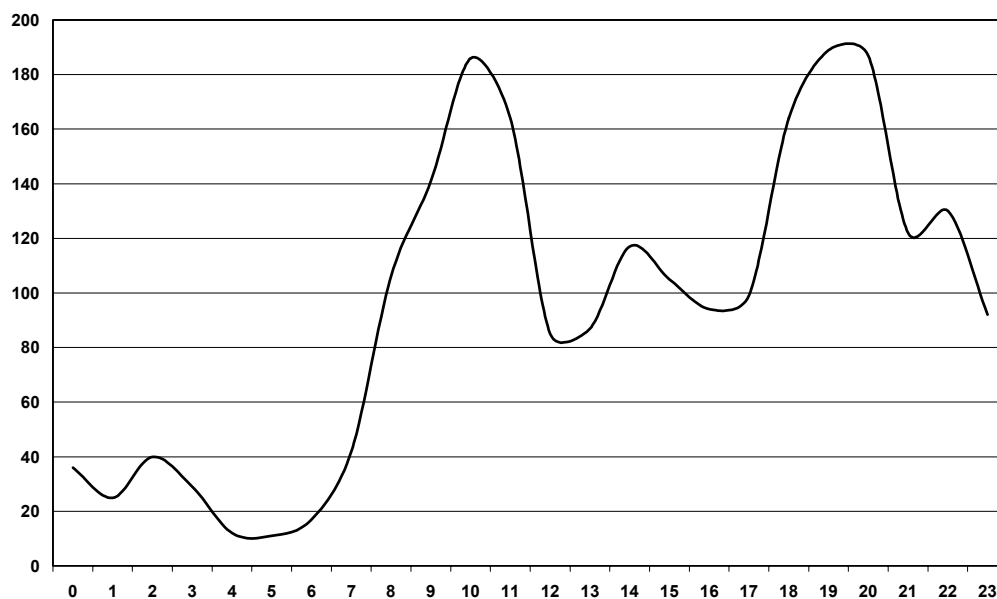
All further analysis was restricted to those 2778 patients who indicated that they *had* spoken with NHS Direct in the previous four weeks. The study populations are shown diagrammatically in Figure 6-3, below.

**Figure 6-3: Study populations**



#### 6.4.1.2 Patients' accounts of their call

The time distribution of calls, by hour, is shown in Figure 6-4, below. This distribution is broadly similar to NHS Direct calls as a whole, though with a greater morning peak than might be expected.<sup>5</sup>

**Figure 6-4: Time distribution of analysed calls**

The advice that respondents reported having been given by NHS Direct is shown in Table 6-6. This distribution of outcomes accords closely with published data from the NHS Clinical Assessment System.

**Table 6-6: NHS Direct triage advice reported by patient**

Advice given	n	%
Contact own doctor	945	34
Self-care advice	914	33
Go to A&E department	411	15
Contact emergency GP	406	15
Put through to 999	93	3
Go to minor injury unit or walk-in centre	64	2
Contact pharmacist	63	2
Contact dentist	54	2
Not sure	30	1
<b>All respondents</b>	<b>2778</b>	<b>100</b>

*Note: a respondent may report being given more than one kind of advice*

A substantial minority of respondents reported being given advice to contact more than one kind or service, and in addition many respondents indicated that they had been given self-care advice as well as advice to contact a service. In order to simplify matters, therefore (which becomes important later in the analysis), where possible we recoded the advice reported by the patient to reflect a single unambiguous course of action recommended by NHS Direct. Where self-care advice was reported alongside advice to contact another service, we took the latter to be the definitive advice. Where only self-care advice was reported, we took this to be the definitive advice. When the patient reported advice to contact more than one service, we took the advice to be, for

our purposes, “ambiguous” (though it should be noted that this does not necessarily mean that the advice was ambiguous for the patient: NHS Direct will often issue advice of the form “do X, but if Y happens, do Z”). However, if the patient indicated they were diverted to 999 we took this as the definitive advice, irrespective of any other advice reported.

In a small number of cases, the patient reported that they had not actually spoken to an NHS Direct nurse for advice, usually because they had waited for a call back and had taken some other action in the interim. The results of the recoding are shown in Table 6-7, below.

**Table 6-7: NHS Direct triage advice, after recoding**

Advice given	n	%
General practice (in or out of hours)	1209	43.5
Self-care only	716	25.8
A&E (with or without ambulance)	480	17.3
Dentist	49	1.8
MIU or walk-in centre	46	1.7
Pharmacy	38	1.4
<i>Ambiguous (multiple advice)*</i>	42	1.5
<i>No advice received*</i>	57	2.1
<i>Question not answered*</i>	141	5.1
<b>All respondents</b>	<b>2778</b>	<b>100</b>

\* These cases are omitted in analyses involving method B compliance, below

Table 6-8 shows, for those advised to contact another service, the time period within which they were advised to do so. Of those advised to contact another service, almost one half of callers were advised to do so immediately, and a further quarter the same day.

**Table 6-8: How soon NHS Direct advised contacting another service, reported by patient**

	n	%
Immediately	747	47.0
The same day	422	26.5
The next day	338	21.3
Not sure	83	5.2
<b>Total</b>	<b>1590</b>	<b>100</b>

In previous work we have explored the issue of compliance with NHS Direct advice, and reasons why patients may be unwilling or unable to follow the advice they are offered.<sup>6</sup> In our 1998 caller survey of first wave NHS Direct sites, we found that 85% of callers said they complied with all of the advice, 13% with some of it and 2% with none

of it.<sup>7</sup> These figures accord closely with patient-reported compliance in the current study (Table 6-9).

**Table 6-9: Whether the patient acted on the advice (Method A: self-report)**

	n	%
All of it	2011	82.3
Some of it	335	13.7
None of it	98	4.0
<b>Total</b>	<b>2444</b>	<b>100</b>

However, a patient's response to the simple question "did you act on the advice?" (method A) may not adequately represent whether they did in fact attend the level of service which was advised. Examination of individual cases revealed that in some instances patients reported that they had not complied with advice, while the contacts they reported were entirely in accordance with the advice they had been given. In some other cases patients reported that they had complied, but their accounts of subsequent events appeared to contradict this. Such apparent inconsistency may be the result of the patient focusing on elements of the advice other than which service to contact, from misunderstanding of our survey, inattention, simple error, or other causes.

In view of this, we compared the patient's account of their actions with the (unambiguous) advice they reported having been given by NHS Direct (method B). In all, unambiguous advice could be determined in 2538 cases (see Table 6-7), of whom 20 did not report the service contacted. Therefore, this measure of compliance could be derived for only 2518 cases.

In each of these cases, we coded compliance as whether the patient did "as advised", did "more" (contacted a service higher in our ranking) or did "less" (contacted a service lower in our ranking) than advised. The results are shown in Table 6-10 below.

**Table 6-10: Whether the patient acted on the advice (Method B: derived)**

	n	%
Did as advised	1781	70.7
Did more than advised	192	7.6
Did less than advised	545	21.6
<b>Total</b>	<b>2518</b>	<b>100</b>

We would argue that this method gives a more accurate picture of what happened, for our purposes, than does the simple self-report described above. It may be noted that the proportion who appear to have fully complied with advice is lower than the proportion who report that they have done so (70.7% v 82.3%).

The great majority of patients (almost 95%) found the advice very or quite easy to follow, as Table 6-11 shows.

**Table 6-11: How easy it was to follow the advice in practice**

	n	%
Very easy	1946	78.7
Quite easy	398	16.1
Quite difficult	79	3.2
Very difficult	49	2.0
<b>Total</b>	<b>2472</b>	<b>100</b>

#### 6.4.1.3 Patients' assessment of NHS Direct advice

The survey asked respondents to rate the appropriateness and helpfulness of the advice they had received from NHS Direct (which might include both triage advice and self-care advice). It is clear that from the point of view of the patient, NHS Direct advice is frequently “very” or “quite” appropriate, although a small proportion – between 7 and 8% – felt it had not been appropriate (Table 6-12).

On the matter of NHS Direct's assessment of the urgency of their problem, there is again a large majority which judged this to have been about right, and between 3 and 5% who felt the degree of urgency was greater or less than assessed (Table 6-13).

**Table 6-12: Patients' views of appropriateness of advice**

	n	%
Very appropriate	1889	70.4
Quite appropriate	593	22.1
Not very appropriate	112	4.2
Not appropriate at all	90	3.4
<b>Total</b>	<b>2684</b>	<b>100</b>

**Table 6-13: Patient's views of appropriateness of timing advised**

	n	%
About right	1462	88.3
Problem was <b>more</b> urgent than NHS Direct said	51	3.1
Problem was <b>less</b> urgent than NHS Direct said	79	4.8
No advice given on timing	64	3.9
<b>Total</b>	<b>1656</b>	<b>100</b>

*Note: this table includes only those who were advised to contact another service*

As might be expected, patients' assessments of the helpfulness of NHS Direct advice show an almost identical distribution to the assessment of appropriateness (Table 6-14).

**Table 6-14: Patients' views of helpfulness of advice**

	N	%
Very helpful	1841	68.7
Quite helpful	600	22.4
Not very helpful	127	4.7
Not helpful at all	113	4.2
<b>Total</b>	<b>2681</b>	<b>100</b>

#### 6.4.2 Classification of necessariness

##### 6.4.2.1 *Contacts with services following the call*

Of the 2778 respondents who had contacted NHS Direct, 60% (1672/2778) reported going on to contact further services, and 40% (1106/2778) did not. Of the 1672 who did contact a further service, 30 did not state the service which they attended. The services attended by the remaining 1642 patients are shown in Table 6-15, below. Almost 90% of those who do contact a further service turn first to general practice or the accident and emergency department. Only about one in four of those contacting a service go on to contact a second service.

**Table 6-15: Contacts with services after the call to NHS Direct**

	First contact		Second contact		Third contact	
	n	%	n	%	n	%
GP or emergency doctor	1148	69.9	214	51.7	31	50.8
A&E department (incl. via ambulance)	313	19.1	79	19.1	6	9.8
Other	51	3.1	16	3.9	5	8.2
Pharmacist	37	2.3	12	2.9	1	1.6
Dentist	34	2.1	5	1.2	1	1.6
Minor injury unit/walk-in centre	30	1.8	7	1.7	0	0.0
Hospital clinic	21	1.3	44	10.6	6	9.8
Helpline	6	0.4	6	1.4	1	1.6
Hospital admission	2	0.1	30	7.2	10	16.4
Physiotherapist/other PAM	0	0.0	1	0.2	0	0.0
<b>Total</b>	<b>1642</b>	<b>100</b>	<b>414</b>	<b>100</b>	<b>61</b>	<b>100</b>

##### 6.4.2.2 *Processes of classification*

Of the 1642 cases presented to our process-based computer program, 1039 (63%) were unambiguously classified on the basis of events reported by the respondent. The remaining 603 (37%) cases were each presented to two members of the panel for review and classification.

Each of these 603 cases was considered by two panel members working independently. In 430 (71%) of the 603 cases, both panel members agreed upon a

classification without discussion. In a further 121 cases, the classification was agreed following discussion between the two raters.

The remaining 52 cases were brought to the full panel for discussion, either because the raters found it difficult to agree or because they felt rating the case might set a precedent which could be generalised to the rating of other cases.

Thus, of the 2748 cases which were potentially classifiable using this method, only 603 (22%) required review by any human rater, and 52 (2%) required review by a panel.

In general, the panel discussions were very straightforward. The panel members found it easy to absorb the principles of the rule-based method, and in most cases found it easy and clinically intuitive to apply the rules to cases in practice. In some cases, however, the panel struggled to apply this approach. The following issues came up:

- In some cases there was simply insufficient information available from the patient survey either to apply a rule unambiguously, or for clinicians to feel comfortable making their classification; without extra information, some events could be interpreted in divergent ways, and it was difficult to avoid “surmising” about the case;
- In some cases the panel found it hard to accept the “rightness” of the processes of care described by the patient, and therefore wanted to judge such cases on the basis of the problem presented rather than the events which actually occurred (but seemed “wrong” to the panel), or on the basis of events which they felt “should” have occurred; for the sake of consistent application of the method, however, the panel was persuaded to classify simply on the events;
- In some cases the panel felt that a problem-based rule, rather than a process-based rule, was the only way forward; these cases provoked the most intense debates. For example, the panel felt that “ill children” under one could not be adequately assessed at nurse-led services, yet there was no obvious way to translate this judgement into a process-based rule. In the end it was simply added as a patient-based rule.

However, despite these difficulties one remarkable feature of this process-based approach was the rapidity and ease with which most cases could be classified in a way which seemed to the panel members intuitive and clinically reasonable.

#### 6.4.3 Necessariness of contacts and advice

The results of the classification of cases are reported in the sections below. It should be noted that the focus of the classification is on the *action taken* by the patient, which may or may not correspond to the *advice offered* by NHS Direct. Thus, we begin by

reporting the action taken. Following this, we restrict the analysis to cases in which the patient appears to have followed the advice, so that we are in effect classifying not simply the actions taken but, by implication, the advice offered to the patient. Finally, we examine the results in terms of both advice offered and action taken in relation to the advice.

#### 6.4.3.1 Classification of actions taken

The results of classifying the necessariness of the action taken by the patient, for all 2748 cases, are shown in Table 6-16 below. These classifications were made under the assumption that all relevant services were available locally, including nurse-led services (walk in centres and minor injury units) with nurse prescribing.

The similarity between the sites is striking. Given that all sites in the study used the same triage software (the NHS Clinical Assessment System), and thus might be expected to produce similar decisions given a similar casemix, this similarity in outcome is what one might expect. The fact that this does indeed occur gives additional confidence in the method used. (In addition, results for NHS Direct in Wales, derived from a study using the same method, are also very similar to those reported here.<sup>30</sup>)

**Table 6-16: Necessariness of first action taken, assuming full service availability**

	Number of cases	Necessary and sufficient (%)	Unnecessary (%)	Necessary but insufficient (%)	Unclassifiable (%)
All cases	2748	81.2	17.1	1.2	0.5
<b>NHS Direct sites</b>					
Site 1	854	81.7	15.7	2.0	0.6
Site 2	881	80.7	18.3	0.8	0.2
Site 3	1013	81.2	17.2	0.9	0.7
<b>Gender of patient</b>					
Male	973	82.5	15.7	1.2	0.5
Female	1768	80.6	17.7	1.2	0.5
Not given	7	57.1	42.9	0.0	0.0
<b>Age group</b>					
0: 9	920	81.2	17.5	0.8	0.5
10: 19	204	79.9	17.2	2.0	1.0
20: 29	360	81.9	16.4	1.4	0.3
30: 39	410	79.8	18.3	1.7	0.2
40: 49	250	85.6	12.4	1.6	0.4
50: 59	214	78.5	19.2	0.9	1.4
60: 69	170	81.2	16.5	1.8	0.6
70: 79	109	82.6	17.4	0.0	0.0
80 and over	61	78.7	19.7	1.6	0.0
Not given	50	84.0	16.0	0.0	0.0

<b>Time of call*</b>					
In hours	554	82.1	16.4	1.3	0.2
Out of hours	1313	80.4	17.7	1.3	0.7
Not given	881	82.0	16.6	1.0	0.5

\* We have taken calls made between 8am and 6pm on Monday to Friday as "in hours" and all other calls as "out of hours". The seven calls made on a bank holiday have been classified as in hours.

The figures in Table 6-16 suggest that, under an assumption that all services are available, the patient's actions following their call proved necessary and sufficient in about four-fifths of cases. In about one in six cases their actions proved to be unnecessary, and in about one case in a hundred their actions proved insufficient.

It should be noted that differences between the various age and call groups may be due to differences in casemix or to differences in triage performance.

Our approach up to this point has assumed that all services in the ranking are available to all patients. However, we have noted above that this is not the case (see Table 6-4) since no walk-in centres are available in one of the areas examined. In general, nurse-led services tend to be fewer in number than other services and therefore may be too far from the patient to be readily accessible. Table 6-17, below, presents the results for calls from a single site in which the classification process was undertaken three times, under different assumptions about the availability of nurse-led services.

**Table 6-17: Necessariness of first action taken, varying assumption of service availability (site 2 only)**

	Number of cases	Necessary and sufficient (%)	Unnecessary (%)	Necessary but insufficient (%)	Unclassifiable (%)
Assuming all services available	881	80.7	18.3	0.8	0.2
Assuming all services available, apart from nurse prescribing	881	83.0	16.0	0.8	0.2
Assuming all services available, apart from any nurse-led services	881	87.6	11.4	0.8	0.2

Under more restricted assumptions about nurse-led service availability, the proportion of patient actions classified as necessary and sufficient rises to close to 90%. This occurs because many patients visit their general practitioner with conditions which could be managed in a nurse-led service, if it were available. Such patients are classified as unnecessary GP attenders under the assumption of full service availability, but become necessary GP attenders if nurse-led services are not available. Under the most restricted service scenario considered here, one in ten patients take actions which prove to be unnecessary.

### 6.4.3.2 Assessment of advice given by NHS Direct

In order to assess the *advice given*, rather than simply the *actions taken*, we now examine the classification of calls after restricting calls only to those in which the patient appeared to have complied with the advice given (using method B, for reasons discussed above). These results are shown in Table 6-18, below.

**Table 6-18: Necessariness of first action taken, compliers only (method B)**

	Number of cases	Necessary and sufficient (%)	Unnecessary (%)	Necessary but insufficient (%)	Unclassifiable (%)
All cases	1781	78.6	20.0	1.0	0.4
<b>NHS Direct sites</b>					
Site 1	556	81.1	16.7	1.6	0.5
Site 2	574	76.5	22.6	0.5	0.3
Site 3	651	78.2	20.6	0.8	0.5
<b>Services available*</b>					
Assuming all services available	574	76.5	22.6	0.5	0.3
Assuming all services available, apart from nurse prescribing	574	79.6	19.5	0.5	0.3
Assuming all services available, apart from any nurse-led services	574	85.7	13.4	0.5	0.3

\* Classification under three service scenarios was undertaken for one site only

Comparing the results above with those in Table 6-16 and Table 6-17, it is apparent that restricting the analysis to full compliers makes only a small difference to the proportions classified as necessary, unnecessary, or insufficient. If anything, the proportion of unnecessary actions is higher among full compliers than among callers as a whole. One might note, however, that selecting calls on the basis of full compliance by the caller does tend to select also those calls in which the advice received seems “right” to the caller, and this may result in a selection bias towards “correctly triaged” calls.

To avoid this potential selection effect, in Table 6-19 the analysis is extended by comparing the action advised by NHS Direct with the action taken by the patient, for all 2518 patients who reported advice from NHS Direct which we were able to code unambiguously, and who reported the nature of their first contact.

This table requires careful interpretation. The services which patients are advised to contact by NHS Direct are shown in the first column. The second column shows how many patients reported being advised to contact the given service. In the third column, these patients are subdivided according to what they did next: “as advised” (i.e. they made the contact advised by NHS Direct, or made no contact if self-care was advised); “less than advised” (i.e. they made either no contact, or a contact lower in

our ranking than that advised by NHS Direct); or “more than advised” (i.e. they made a contact higher in our ranking than that advised by NHS Direct).

The figure in the fourth column shows how many patients behaved in the way described. Subsequently, the remaining columns classify, in the same terms as before, the action *actually taken*. Thus, while the classification of action can be taken to classify the advice for those patients who have complied, this is *not* the case for those patients who have not complied. Nonetheless, even for non-compliers the results may be informative, as we discuss below.

**Table 6-19: Necessariness of first action taken, by advice and action**

Action advised	n	Action taken	n	Necessary and sufficient			
				Unnecessary	Necessary but insufficient	Unclassifiable	
A&E	471	Did as advised	266	131	135	0	0
		Did less than advised	205	188	12	3	2
GP	1200	Did as advised	880	639	218	15	8
		Did more than advised	32	14	17	0	1
		Did less than advised	288	282	3	3	0
Dentist	49	Did as advised	23	22	1	0	0
		Did more than advised	5	2	2	1	0
		Did less than advised	21	21	0	0	0
MIU	46	Did as advised	16	15	0	1	0
		Did more than advised	17	12	5	0	0
		Did less than advised	13	13	0	0	0
Pharmacy	38	Did as advised	12	10	2	0	0
		Did more than advised	8	6	1	1	0
		Did less than advised	18	18	0	0	0
Self care	714	Did as advised	584	583	0	1	0
		Did more than advised	130	89	39	1	1
<b>Any advice</b>	<b>2518</b>	<b>Did as advised</b>	<b>1781</b>	<b>1400</b>	<b>356</b>	<b>17</b>	<b>8</b>
		<b>Did more than advised</b>	<b>192</b>	<b>123</b>	<b>64</b>	<b>3</b>	<b>2</b>
		<b>Did less than advised</b>	<b>545</b>	<b>522</b>	<b>15</b>	<b>6</b>	<b>2</b>

For each service to which NHS Direct triages, the table above gives some indication of the “predictive value” of the triage advice – that is, the likelihood that contact with the service will prove to be both necessary and sufficient. For example, of those who comply with advice to attend A&E, approximately half (131/266) will prove to have a necessary attendance and half will not; and of those who comply with advice to contact a GP, approximately three-quarters (639/880) will prove to have a necessary and sufficient attendance. Such results might be regarded as acceptable given that one of the challenges which faces NHS Direct is to identify the serious, urgent cases among the very large number of non-serious, non-urgent problems which present.

However, the table also indicates that frequently patients choose not to comply with NHS Direct, and that in such cases patients may still take action which proves to be necessary and sufficient. Taking all the cases together, the results indicate that callers who act on NHS Direct advice are less likely to make unnecessary contacts (356/1781, 20%) than those who do more than advised (123/192, 33%), as might be expected. Those who do less than NHS Direct advises are far less likely to make an unnecessary contact than either of these groups (15/545, 3%), though they face a small additional risk of doing too little.

Of course, one should remember that the casemix of each of these groups –compliers and non-compliers – may be different, so that they are not strictly comparable. The table reminds us that NHS Direct is not a technology which is simply “applied” to callers, but is a social interaction in which nurse advice is interpreted by the patient in the light of their existing knowledge and experience, their current circumstances, and the degree of trust they are willing to place in the nurse. In this sense the outcome of an NHS Direct consultation is a “co-production” of the caller, the nurse and the computer.<sup>31</sup>

From this perspective, one might suggest that callers treat the advice they receive from NHS Direct simply as further information to “take into account”, just as a clinician might do with the results of a diagnostic test. In Bayesian terms, callers have a “prior belief” about the necessity of seeking health care, which is then modified by the advice which NHS Direct offers, and the way in which it is offered (since nurses may express the advice forcefully or tentatively). Callers with a strong prior belief, in either direction, will be less likely to have this altered by NHS Direct advice than those with a weaker prior belief; and conversely, NHS Direct advice which is delivered with greater certainty or force is more likely to alter a caller’s prior belief than advice which is delivered uncertainly. Overall, the combination of the caller’s prior belief with the advice from NHS Direct may well produce better (more necessary and sufficient) outcomes than those which would result from simple compliance with all advice.

Thus, we can distinguish two levels on which we can assess the quality of NHS Direct advice. At the pragmatic level, we are interested only in what actions are taken by the patient after the call, irrespective of whether the patient has complied with the advice. On this level, we are not concerned with whether the advice given was “right” or “wrong”, but only with the outcome (the action taken). This is the approach described in the previous section, which leads us to conclude that approximately 80% of calls results in necessary and sufficient action, 17% in unnecessary action and about 1% in insufficient action.

However, we may wish also to assess NHS Direct advice at the explanatory level: i.e. how good was the advice given? This is a harder question, both to answer and for NHS Direct to perform well on. However, the data presented in Table 6-19, which considers outcome against compliance, does give an indication of performance. Of

the 2518 calls assessed here, in 1400 the caller did as advised and their action was rated “necessary and sufficient”; in 64 they did more than advised and their action was rated “unnecessary” and in 6 they did less than advised and their action was rated “insufficient”. Thus, in 1470/2518 (58%) of calls one might argue the advice was broadly correct.

Similarly, 356 callers acted on NHS Direct advice and their action proved unnecessary; 522 did less than advised, and this proved sufficient; and 15 did less than advised and this proved unnecessary. Thus, in 893/2518 (35%) of calls, NHS Direct appears to be over-triaging. Finally, 17 callers acted on advice and were rated insufficient; 123 did more than advised and this proved to be necessary; and 3 did more than advised and this still proved insufficient. Thus, in 143/2518 (5.7%) of calls, NHS Direct appears to be under-triaging.

It should be noted, however, that this further analysis depends upon deriving a simple compliance variable which may not adequately represent the complexities of the call. For example, NHS Direct may frequently advise callers “if it gets worse, do X”. We have not attempted to capture this kind of advice in the current study, so it is possible that our analysis of advice and compliance misrepresents the true picture. Nonetheless, we have included it here because it gives an indication of how this method could be developed in future studies.

#### 6.4.4 Comparison of self-rating and process-based rating

We have noted above that over 90% of patients regarded the advice they had been given by NHS Direct as “very” or “quite” appropriate (Table 6-12). In Table 6-20 below we compare the views of patients on the appropriateness of the advice they received with the rating we assigned using our process-based approach. It may be seen that, for calls in which the patient is unhappy with the advice given, the proportion of calls with actions rated as unnecessary or insufficient in the process-based classification rises. This suggests some degree of concordance between the patient’s view and the process-based measure.

**Table 6-20: Necessariness of first contact by patient’s view of appropriateness of advice given**

Patient's assessment	Number of cases	Necessary and sufficient (%)	Unnecessary (%)	Necessary but insufficient (%)	Unclassifiable (%)
Very appropriate	1868	83.0	15.6	0.9	0.4
Quite appropriate	588	78.9	19.7	0.9	0.5
Not very appropriate/ Not appropriate at all	199	70.9	22.1	5.5	1.5
<b>Total</b>	<b>2655</b>	<b>81.2</b>	<b>17.0</b>	<b>1.2</b>	<b>0.5</b>

Of course, it may be argued that in calls where they are unhappy, patients are less likely to comply with advice and may take unadvised – and indeed ill-advised – actions which result in unnecessary or insufficient contacts. To examine this possibility, we again restricted the analysis to those reporting full compliance, so that the actions taken reflect the triage advice given, and these results are shown in Table 6-21, below. The relationship between the caller’s view and the process-based rating is again clearly evident.

**Table 6-21: Necessariness of first contact by patient’s view of appropriateness of advice given, compliers only (Method A)**

Patient's assessment	Number of cases	Necessary and sufficient (%)	Unnecessary (%)	Necessary but insufficient (%)	Unclassifiable (%)
Very appropriate	1514	80.6	18.0	0.9	0.4
Quite appropriate	382	75.4	23.0	0.8	0.8
Not very appropriate/ Not appropriate at all	77	63.6	31.2	5.2	0.0
<b>Total</b>	<b>1973</b>	<b>79.0</b>	<b>19.5</b>	<b>1.1</b>	<b>0.5</b>

## 6.5 Conclusions

In the study reported here we have developed a method of assessing the appropriateness of NHS Direct triage advice in terms of the necessariness of the service contact which follows, and have applied it to 2748 calls to NHS Direct. In summary, the results suggest that in 76 to 86% of cases (depending on assumptions made about service availability) NHS Direct gives advice which results in necessary and sufficient use of subsequent services. In 11 to 22% of cases it gives advice which results in arguably unnecessary use of services, and in about 1% of cases it gives advice which results in use of services insufficient to meet the patient’s needs.

Taking these results at face value, we would note that this pattern of results is consistent with other evidence suggesting that NHS Direct tends to err on the side of caution in triaging patients.<sup>6</sup> Clearly, a degree of risk aversion in a service such as this is a desirable characteristic, since it is likely to produce a service which is safer for patients than might otherwise be the case. Against this, one must balance the inconvenience to the patient and costs to the service and to other patients (in terms of increased waiting times) of sending a patient unnecessarily to a service they do not need. NHS Direct nurses face the difficulty of balancing these risks in the triage decisions they make. Whether the balance is currently the “correct” one cannot be determined from a study such as this, since it is in essence a social judgement.

However, the results we have reported do suggest that there may be a substantial minority of calls in which NHS Direct misjudges the severity of the symptoms being reported and over-triages the patient. The results also suggest that in some of these calls this may be evident to the patient, either at the time of the call or later (since

there is an association in the data between apparent mistriage, lower compliance and lower patient rating of appropriateness). In such cases, it may be that there are remediable factors such as over-cautious decision support algorithms which, if identified and corrected, might reduce the proportion of over-triaged calls without increasing the proportion which are under-triaged. Thus, application of the method we have described could prove to be valuable both in targeting clinical audit efforts in NHS Direct towards potentially incorrect triages, and in measuring the effectiveness of any measures taken as a result. Whether in fact telephone triage can be improved in ways which increase the specificity of triage without harming sensitivity remains an open and important question.

In the current study, a random selection of apparently mistriaged calls is to be examined in further detail by the clinical panel in order to determine whether calls which we have rated as over or under-triaged in terms of the processes of care required were nonetheless justifiable or at least explicable in other terms. The results of this will be reported in due course.

#### 6.5.1 Strengths and limitations of this study

This study has a range of important strengths and weaknesses. The major strength of the approach we have described is that the idea of “appropriateness” which, as we have seen, is both conceptually and operationally difficult or impossible to define, has been translated into the idea of “necessariness”. We have shown how this idea can, in practice, be defined in a way which makes sense to clinicians and can be operationalised as a set of explicit and simple rules. We have also shown that this operational definition can be applied across a large number of calls.

Thus, the current study may be viewed as a “proof of concept” of the process-based approach in this area. It might be objected that we have not attempted to “validate” the classification outcomes of this approach, and this is true. This is because “validation”, for example by reference to an “expert opinion” on whether each contact was in fact necessary, would simply return us to the situation of being reliant upon subjective and implicit judgements of appropriateness. Instead, it is better to see the process-based method as straightforwardly identifying cases which do or do not conform to a set of explicit rules. The method should then be assessed in terms of whether or not the rules make sense, rather than in terms of the results of applying them, bearing in mind that it is in principle possible to reclassify cases under a different set of rules if necessary.

Since the purpose of the panel is also primarily in terms of interpreting and applying process-based rules, it is likely that the current approach is less sensitive to the composition of the panel than previous attempts to measure appropriateness.

### 6.5.1.1 Methodological limitations

There are a number of limitations inherent in the method we have described. To begin with, the approach requires a number of basic assumptions to be made (Section 6.2.3 above), each of which may be challenged. For example, there were difficulties in practice in defining a “ranking of care” (assumption one). There were arguments for and against placing nurse-led services below general practice, and there was no clear place in the ranking for dentists. In principle, it would be possible to develop the method further using a more complex ranking in which services do not fall in a simple linear order in relation to one another, or in which the order varies for different kinds of problem (for example, a ranking for one clinical condition, such as trauma, might differ from the ranking for other conditions, such as non-trauma). However, some idea of “higher” and “lower” is intrinsic to our approach.

Similarly, in some cases the panel did not feel that the processes of care described by patients indicated that everything necessary had been done, or that unnecessary things had not been done (assumption three). It is of course entirely possible that patients misreport events, or that providers do the wrong things. As a matter of principle, however, we felt it best always to apply the rules to what had actually been done, rather than to what we felt *should* have been done, since the panel cannot know exactly how the problem presented to the clinician, or what circumstances or negotiations had to be taken into account at the time.

It may be objected that providers may consistently do things which are unnecessary, just because they can. For example, A&E doctors may X-ray conditions which would be managed adequately without X-ray in general practice. If this bias were operating, then we would have *overestimated* the proportion of contacts which were necessary.

However, against this we would argue that the decision to X-ray depends on the provider’s perception of risk, and such risk perceptions (or Bayesian priors) are bound to be different in A&E than in general practice. Thus, the increased use of X-ray in A&E may be clinically appropriate to the increased likelihood of serious conditions, rather than simply an expression of availability.

As we have noted previously, the fact that we rely upon the patient’s account may mean that some things which occur, but are unobservable to the patient, cannot be reported. In particular, if a GP examines a patient by ‘general observation’ but in no other way, this will not be apparent to the patient and is likely to go unreported, so that we will incorrectly regard the contact as unnecessary.

A further limitation of the process-based method is that, while it assesses the *level*, it is unable to assess the *timing* of the contact. Thus, in the current study we have not distinguished between patients attending the A&E department by ambulance or by other transport, or between patients seeing a GP in hours or out-of-hours. In addition,

the method is unable to assess the advice given in terms of other indicators of quality, such as promptness, clarity, comprehensiveness, and so on.

More generally, making an assessment of necessariness based upon rules about what services are available where is clearly to take an NHS-centric view of the world, in which we wish to see the “rational use” of scarce NHS resources. We take no account of the patient’s view which may be, for example, that additional reassurance is available at A&E which is not available from the GP.

#### *6.5.1.2 Practical difficulties*

There were also various difficulties in carrying out the project which may limit the strength of the findings, although in principle these may be overcome in future applications of the method.

The response rate to the survey was lower than we had hoped, and this naturally introduces a concern over response bias. Although the characteristics of patients were broadly in line with previous studies of NHS Direct, there was evidence that response rates among young adults (15 to 30 year olds), especially, men, were lower than for other groups. In addition, we have noted that an important reason for non-response may be serious illness or death. We have attempted to address this issue through a survey of the GPs of non-responders, though this has proved difficult to arrange and is still in progress.

In addition, we rely entirely on the patient’s report of the events which occurred. Although it would be possible, and desirable in terms of case processing, to ask more questions about the processes of care, the survey is already long and adding further questions may lower response or item completion rates. However, it might be possible in future versions of the survey to improve the design and layout in ways which aid clarity and add detail while not proving off-putting to patients. If this were possible, it might be possible to improve response rates while increasing the proportion of cases which can be directly processed by computer without reviewer assistance.

Finally, we have not attempted, in this study, to compare the performance of NHS Direct with that of other triaging services, such as GP co-operatives, although such a comparison is, in principle, possible. Indeed, one of the virtues of the method described here is its potential value in comparative studies, particularly for evaluating the effects of interventions. The effects of changes to NHS Direct, such as updates to the Clinical Assessment System, changes in nurse training, or availability of other services, on the quality of advice given and the actions taken by callers, could be rapidly and objectively assessed using this approach.

## **7. CONCLUSIONS**

### **7.1 Summary of findings**

We have reported here on three studies which, directly or indirectly, examine the impact of NHS Direct on the use of other NHS services. Using data from ambulance, A&E and GP co-operative services across England, Wales and Scotland, and examining the effect of the introduction of each successive wave of NHS Direct, we have concluded that NHS Direct has been associated with a fall in the number of calls to co-operatives, but with no important change in demand for A&E or emergency ambulance services. These findings are consistent with our previous work in this area.<sup>12</sup> Using large population surveys in the areas longest served by NHS Direct, over four successive years, we draw similar conclusions. These data suggest no important change in the use of other services, and no change in the mean number of health care services used during an unplanned episode of care. They also suggest that, although NHS Direct call rates are rising, the service is used in only a small proportion (6% or less) of unplanned episodes of care.

In the third study we have examined the appropriateness of NHS Direct triage advice in 2748 calls in terms of the “necessariness” of contacts with services which follow the call. Under the most realistic assumption, that nurse-led services are unavailable locally, this study suggests that advice leads to necessary contacts in 88% of calls, unnecessary contacts in 11% of calls, and insufficient contacts in about 1% of calls. Further work is underway to determine whether the reasons for apparent over-triage are evident during the call.

### **7.2 Implications for further research**

Now that the whole of England and Wales, and parts of Scotland, are served by NHS Direct or NHS24 quantitative research into the effects of the service on overall demand for health care is becoming difficult or impossible, since there are few remaining areas with which comparisons can be made. However, since the service is now an established part of NHS immediate care services, which are beginning to reconfigure around it, it could be argued that the question is no longer as important as others, such as those relating to the quality of the service and the advice offered.

The assessment of the quality of NHS Direct triage and self-care advice remains a key issue for further research, as the service takes on new roles (such as triaging for co-operatives and ambulance services) and the decision support systems are developed further. The process-based method used in this study has potential to be developed further to overcome some of the limitations we have noted above, and could be applied in other NHS Direct sites on a periodic basis to assess the success of the service as it develops. In addition, a similar approach could be developed for specific conditions of interest. Practical issues which need to be addressed include improving

the data collection phase so that response and completion rates are higher, surveying is faster and data entry is fast and reliable.

In addition, qualitative work in this area is now needed to understand the important factors at work in calls which result in unnecessary or insufficient use of health services.

### **7.3 Implications for policy**

As NHS Direct has developed and integrated with other services, the issue of the quality of the triage and self-care advice offered has become a central concern to other clinical professionals. In order to remain credible, NHS Direct triage must be open to explicit and reproducible assessment which can demonstrate how the service is performing and how it is changing in response to improvements in management, training, decision support systems and so on. Although clinical audit of NHS Direct calls has been carried out since its inception, this has tended to focus on within-call processes and has infrequently examined the contribution of the service to the patient's entire episode of care. Where the patient's care after the call has been examined, there have been difficulties in reaching agreement on what is "appropriate".<sup>26</sup> The process-based approach to assessing triage advice we have described may therefore prove to be useful in widening the scope of NHS Direct's audit and bringing some objectivity to this area.

The results we have reported here suggest that NHS Direct triages to necessary and sufficient services in a large majority of cases. There is also evidence, however, that about one in eight callers may receive advice leading to an unnecessary contact with health services. The challenge for NHS Direct now is to reduce this proportion, while at the same time maintaining the safety of the service. We do not yet know whether this can be achieved.

## **8. ACKNOWLEDGEMENTS**

We are very grateful to the members of the clinical panel who gave so much of their time and effort to this study: Ken MacDonald, Suzanne Mason, Helen Metcalf, Fizz Thompson and Mike Vaughan. In addition, we are grateful to the very many staff of GP co-ops, ambulance services and emergency departments nationwide who contributed data to the impact study, to the staff of agencies providing population samples for our surveys, to the Royal Pharmaceutical Society of Great Britain for providing the list of drugs, and to the staff of NHS Direct who gave valuable help in undertaking patient surveys.

This study was funded by the Department of Health. The views expressed are those of the authors and not necessarily those of the Department of Health.

## 9. REFERENCES

- 1 Department of Health. *The New NHS: modern, dependable*. Cmd 3807. London: The Stationery Office, 1997.
- 2 Calman K. *Developing emergency services in the community: the final report*. London: NHS Executive, September 1997.
- 3 Scottish Executive. *It's NHS 24 – Deacon announces details of new 'all day, every day' NHS phone service*. Press release SE3208/2000. Edinburgh: Scottish executive, 13 December 2000.
- 4 Munro J, Nicholl JP, O'Cathain A, Knowles E. *Evaluation of NHS Direct first wave sites: first interim report to the Department of Health*. Sheffield: Medical Care Research Unit, 1998.
- 5 Munro J, Nicholl JP, O'Cathain A, Knowles E. *Evaluation of NHS Direct first wave sites: second interim report to the Department of Health*. Sheffield: Medical Care Research Unit, 2000.
- 6 Munro J, Nicholl JP, O'Cathain A, Knowles E, Morgan A. *Evaluation of NHS Direct first wave sites: final report of the phase 1 research*. Sheffield: Medical Care Research Unit, 2001.
- 7 O'Cathain A, Munro JF, Knowles E, Nicholl JP. How helpful is NHS Direct? Postal survey of callers. *BMJ* 2000;**320**:1035.
- 8 Boseley S. Just hanging on the telephone. *The Guardian*, 9 January 2001.
- 9 Hellen N, Carr-Brown J. NHS phone lines 'put lives at risk'. *Sunday Times*, 3 December 2000.
- 10 Williams S. NHS Direct investigated. *Health Which?*, August 2000: 12-16.
- 11 Mayor M. Health watchdog criticises NHS helpline. *BMJ* 2000;**321**:401.
- 12 Munro JF, Nicholl JP, O'Cathain A, Knowles E. Impact of NHS Direct on demand for immediate care: observational study. *BMJ* 2000;**321**:150-3.
- 13 Chapman RS, Smith GE, Warburton F, Mayon-White RT, Fleming DM. Impact of NHS Direct on general practice consultations during the winter of 1999-2000; analysis of routinely collected data. *BMJ* 2002;**325**:1397-8.
- 14 O'Cathain A, Webber E, Nicholl J, Munro J, Knowles E. NHS Direct: consistency of triage outcomes? *Emergency Medicine Journal* 2003;**20**:289-92.
- 15 Jones Elwyn G, Stott NCH. Avoidable referrals? Analysis of 170 consecutive referrals to secondary care. *BMJ* 1994;**309**:576-8.
- 16 Khan KL, Kosecoff J, Chassin MR et al. Measuring the clinical appropriateness of the use of a procedure: can we do it? *Medical Care* 1988;**26**:415-22.
- 17 Working Group for Director of R&D of the NHS Management Executive. What do we mean by appropriate care? *Quality in Health Care* 1993;**2**:117-123.
- 18 Hicks NR. Some observations on attempts to measure appropriateness of care. *BMJ*

- 
- 1994;**309**:730-3.
- 19 Brook RH, Kosecoff JB, Chassin MR, Winslow CM, Hampton JR. Diagnosis and treatment of coronary heart disease: comparison of doctors' attitudes in the USA and UK. *Lancet* 1988;**i**:750-3.
- 20 Buetow SA, Sibbald B, Cantrill JA, Halliwell S. Appropriateness in health care: application to prescribing. *Social Science and Medicine* 1997;**45**:261-271.
- 21 O'Cathain A, Sampson F, Nicholl J P, Munro J F. *The effect of skill mix on clinical decision-making in NHS Direct*. Report to West Midlands NHS Executive Director of R&D. Medical Care Research Unit, 2001.
- 22 Grant C, Nicholas R, Moore L, Salisbury C. An observational study comparing quality of care in walk-in centres with general practice and NHS Direct using standardised patients. *BMJ* 2002;**324**:1556-9.
- 23 Cave AJ. Results of study on walk-in centres are only to be expected. *BMJ* 2002;**325**:839.
- 24 Gaffney P, Crane S, Johnson G, Playfair M. An analysis of calls referred to the emergency 999 service by NHS Direct. *Emergency Medicine Journal* 2001;**18**:302-4.
- 25 Bland JM, Altman DG. Some examples of regression towards the mean. *BMJ* 1994;**309**:780.
- 26 BBC News Online. NHS Direct referrals criticised. 17 March 2001. At: <http://news.bbc.co.uk/1/hi/health/1224576.stm>
- 27 Lowy A, Kohler B, Nicholl J. Attendance at accident and emergency departments: unnecessary or inappropriate? *Journal of Public Health Medicine* 1994;**16**:134-140.
- 28 Martin A, Martin C, Martin PB, Martin PAB, Green G, Eldridge S. "Inappropriate" attendance at an accident and emergency department by adults registered in local general practices: how is it related to their use of primary care? *Journal of Health Services Research and Policy* 2002;**7**(3):160-5.
- 29 Royal Pharmaceutical Society of Great Britain. *Medicines, ethics and practice*. (MEP). No 25. London: RPSGB, July 2001.
- 30 Rance, J, *et al*. *Evaluation of NHS Direct in Wales*. Forthcoming.
- 31 O'Cathain A, Sampson FC, Munro JF, Thomas KJ, Nicholl JP. Nurses' views of using computerized decision support software in NHS Direct. *Journal of Advanced Nursing* 2004;**45**:280-6.