Surveillance for Health Protection in England and Wales

An analysis of NHS Direct syndromic data

Duncan L Cooper

ISBN 978-90-9023368-0

http://www/nivel.nl Tel: +31 30 27 29 700
http://www.hpa.org.uk Tel: +44 121 352 5066

Published papers were reprinted with permission of Oxford Journals, Cambridge Journals, BioMed Central, British Journal of General Practice, Health Protection Agency, MMWR Publications and Elsevier Publishing.

Cover design: © Nicola Moorhouse
Printing: PrintPartners Ipskamp, Enschede

Surveillance for Health Protection in England and Wales

An analysis of NHS Direct syndromic data

Surveillance voor Gezondheidsbescherming in Engeland en Wales:

Een analyse van syndroomgegevens van NHS Direct

(met een samenvatting in het Nederlands)

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Utrecht,
op gezag van de rector magnificus, prof. dr. J.C. Stoof,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
donderdag 18 september 2008 des ochtends te 10.30 uur

door

Duncan Lewis Cooper
geboren op 21 november 1968
te Londen, Verenigd Koninkrijk

Promotoren:

Prof. dr. P.P. Groenewegen

Prof. dr. J. van der Zee

Co-promotor:

Dr. G.E Smith

The studies presented within this thesis were conducted by the Health Protection Agency and NHS Direct with contributions from staff of the University of Manchester, Health Protection Scotland, Netherlands Institute for Health Services Research (NIVEL) and Universiteit Utrecht.

Contents

1 General Introduction

Summary

Disease surveillance is the collection and analysis of health data to provide information for action and to inform decisions relating to public health policy. This action may be an acute response such as detection of an infectious disease outbreak, or contact tracing and targeted intervention. Alternatively, the surveillance activity may occur over a longer time period as surveillance data, for example, firstly provide evidence to support a new vaccination policy and then measure its impact over subsequent years. Surveillance systems in the UK typically rely on data about diagnoses made by clinicians, mostly within primary care centres (Fleming, 1999), or data about laboratory confirmations of specific disease pathogens (Hawker et al., 2001). However, in recent years there has been a growth in syndromic surveillance systems that collect and analyse pre-diagnostic data in order to estimate the health status of the community (Henning, 2004): pre-diagnostic meaning before a diagnoses is made. Such systems commonly employ 'real-time' data analysis in an attempt to provide more timely warning of infectious disease outbreaks, epidemics, pandemics or potential bio-terrorist attacks. They often aim to identify an increase in illness before diagnoses are confirmed and reported to public health agencies.

Despite initial enthusiasm and optimism many syndromic surveillance systems have not wholly met their aims, prompting calls for a diversification of data sources and further evaluation of what purpose they can and do serve (Rheingold, 2003). Numerous county, city or state wide systems operate in the US using data about hospital emergency department visits (Heffernan et al., 2004) and ambulance encounters (Lazarus et al., 2001), although primary care clinician visits (Miller et al., 2004) and over the counter pharmacy sales (Das et al., 2005) are also employed. Major issues that those running these systems face are sensitivity (the size of event that can be detected), sustainability, and developing baselines levels of reported syndromes from which to detect unusual events.

Since 2001 the UK Health Protection Agency (formerly the Public Health Laboratory Service) has used daily data from NHS Direct, a *national* UK telephone triage system ('nurse hotline'), for syndromic surveillance purposes. These data may be of particular value because NHS Direct is available to the entire population of England

and Wales, callers report a wide range of conditions, and data are collected in a consistent format. Also, NHS Direct is many peoples first or only point of contact with the health service. Consequently there may be an opportunity to identify an increase in illness before it is reported to other primary care or secondary care services, thus providing the first signs of an emerging public health problem. Finally, a crucial advantage of this system is that the data are already gathered as part of a centrally managed national clinical care system, making it considerably more cost-efficient than systems developed primarily for surveillance.

The capability of disease surveillance to lead to the early identification of a public health problem was demonstrated as early as 1854 when Dr John Snow mapped cases from a cholera outbreak in Central London's Soho district (Snow, 1854). This led him to remove the handle of the water pump he believed to be the source of the disease and ultimately control the outbreak. In the 21st century medical geography is a well established discipline traditionally divided into the study of the distribution of disease (incorporating elements of spatial epidemiology), and the geography of health care. Infectious disease epidemiologists in particular have embraced space-time cluster detection methods (Kulldorff et al., 2005) to gather evidence of unusual disease clusters (Jones et al., 2006) or describe the spatial diffusion of epidemics (Mostashari et al., 2003). As opposed to a priori investigations, real-time surveillance has the potential to proactively provide early warning of disease threats. It may also be possible, using spatio-temporal data analysis, to consider the type of disease diffusion present within the data. For example, Cliff and Haggett describe contagious disease diffusion requiring person to person contact and the spread of disease outwards through a population (Cliff and Haggett, 1981). Viral respiratory pathogens such as influenza and measles are diseases that spread in this way.

Aims

The NHS Direct syndromic surveillance system is the only national surveillance system in the UK that monitors disease incidence in a real-time fashion and to a local level. In this thesis we describe and evaluate the use of NHS Direct call data for surveillance purposes. The primary research question is: "What is contribution of NHS Direct data to health protection surveillance in England and Wales?" Health protection surveillance is the surveillance of infectious diseases and environmental hazards that may pose a threat to public health. We explore this research question with reference to the principles of modern surveillance systems (HPA, 2005);

- that they have clearly defined aims,
- are flexible and representative of the population, and
- achieve an acceptable sensitivity, specificity and positive predictive value.

We also address whether routine analyses of NHS Direct call data can be used to;

- provide early warning of outbreaks of respiratory or gastrointestinal disease,
- detect and quantify significant trends in disease, and
- identify individual cases of public health importance.

Most importantly, we discuss ways in which surveillance of NHS Direct call data informs public health action.

The studies in this thesis demonstrate how timely NHS Direct syndromic surveillance system has been used to serve multiple public health and scientific purposes. We attempt to characterise the diffusion of influenza and norovirus (both common viruses) by studying spatio-temporal changes in syndromic call rates across England. We also present a range of statistical techniques to explore the purely temporal variation in NHS Direct calls. A control chart methodology is proposed for detecting significant rises in NHS Direct calls above critical threshold values. These values vary according to syndrome, geographical area and day of the week. The effects of age, sex and social deprivation on NHS Direct call rates are examined using a negative binomial regression model. *Poisson* regression models are used to derive threshold values for influenza surveillance which are tested for their potential to provide early warning of national rises in influenza during five consecutive winters. Multiple linear regression models are constructed to quantify the seasonal effects of a range of infectious disease pathogens on demand for NHS Direct. Finally, we use statistical simulation to inject a real outbreak of cryptosporidiosis into our surveillance system to determine whether this outbreak would have been detected before traditional approaches.

After the following introductory chapter, the body of this thesis is divided into three parts. Part One (chapters 2-4) introduces the NHS Direct service, and describes the aims, methodology and representativeness of the NHS Direct syndromic surveillance system. Part Two (chapters 5-8) contains analytical studies that compare NHS Direct data against data from other surveillance systems (notably laboratory and clinical) and examines the utility for outbreak detection and providing early warning of rises in disease. Part Three (chapters 9-11) presents results from the syndromic surveillance system and a validity study to determine whether it is feasible to link syndromic surveillance directly to microbiological testing. The conclusions chapter discusses the findings of this thesis with reference to other work in this field and also the implications for public health practice.

Introduction

The remainder of this chapter describes why NHS Direct was introduced in the UK, who uses the service and the factors that influence demand. The concept of disease surveillance is then introduced and placed within the context of the UK Health Protection Agency. Current surveillance arrangements are described with reference to their potential shortfalls and how NHS Direct data may fill some of these gaps. The rationale behind syndromic surveillance is then covered and why NHS Direct data may provide a unique and timely snapshot of community morbidity in England and Wales. Finally, the aims and structure of the thesis are presented.

Telehealth and NHS Direct

The UK National Health Service (NHS)

The UK National Health Service is the publicly funded health care system of the United Kingdom. It was established in 1948 with the overall aim of improving the health and well-being of the population through the provision of high quality health and social care services to all. At the time of writing (March 2008) 10 Strategic Health Authorities (SHAs) form the link between the Department of Health, who fund the NHS, and the NHS itself. SHAs are responsible for the strategic direction of the NHS locally and monitor the performance of Primary Care, Community and Hospital Trusts. The performance of Care Trusts (combining health and social care), Mental Health Trusts and Ambulance Trusts are also the responsibility of SHAs.

Since 2002 Primary Care Trusts (PCTs) have been at the heart of the NHS, controlling and spending 80% of the total budget. They are local organisations (numbering 151 in England in October 2007) reporting to SHAs and have overall responsibility for assessing local health needs, commissioning local services to meet these needs, and improving the health of their local communities. As well as buying and monitoring services (e.g. GP practices, pharmacies, dentists, NHS Direct), PCTs have a developmental role in encouraging NHS organisations to maximise the quality and access to local services.

Telephone triage

Due to the increasing demand for out-of-hours primary care in the UK during the 1980s and 1990s (Salisbury, 2000), alternative approaches to the on-call doctor were considered, including telehealth solutions. Telehealth is the use of electronic information and communication technologies to support remote clinical health care, and patient and professional health-related education (Tizzard and Edwards, 2005). Innovative use of technology to deliver telehealth services, such as by video or audio-conferencing, means that health care services can be provided to those who are some distance from the provider, thus reducing geographic barriers to accessing care. Telehealth applications now cover a diverse range of health care services, including psychiatry, cardiology, gastroenterology, rehabilitation, ultrasound, remote diagnostics, dialysis, and robotic surgery. Telehealth is also becoming increasingly vital to developing nations by providing health care to rural populations, although technological barriers and high costs associated with initial investment may constrain this uptake.

In 1997, the World Health Organisation (WHO) announced that it would make telehealth a global strategy area for the 21st century, advocating its use for disease prevention, education and training, and interestingly, disease surveillance (Schneider, 1998). Since this date various initiatives have been launched to provide free online access to health journals and information for nonprofit organisations, and eLearning initiatives for the public in developing nations. Although WHO still predominantly promotes rather than initiates the use of telehealth for patient care, pilot telehealth projects for patient management are underway in South East Asia (WHO, 2005).

Telenursing or *telephone triage* is a subset of telehealth in which the focus is on nursing practice via telecommunications (usually the telephone). Such systems have gained in prominence as governments and major health care providers have sought to control rising health care costs, reduce inappropriate visits to emergency care, and meet patient demand for round-the-clock health care advice and information. In countries with largely public health care delivery, telephone triage systems now operate on regional (e.g. Telehealth Ontario in Canada, HealthDirect in Western Australia) and national levels (e.g. NHS Direct in England and Wales, NHS 24 in Scotland, Healthline in New Zealand). In other countries, such as the Netherlands, a number of health insurance organizations have set up nurse led information lines for their insured. Privately run telephone triage systems managing those with health insurance have become commonplace with an estimated 100 million Americans able to access some form of telephone triage (Hutcherson, 2001).

NHS Direct (1998)

In the 1980s, nurse triage or doctor consultation via the telephone began to be introduced by providers of out-of-hours care in the UK and was found (particularly when clinical decision support software was used) to be as safe and effective as face to face care (Lattimer et al., 1998). Its success, coupled with the growing demand for out-of-hours health care and advice, led to the announcement of a new primary care service - NHS Direct - in the 1997 UK Government White Paper 'The New NHS: Modern and Dependable' (DH, 1997). NHS Direct was to be a nurse-led health helpline providing the population of England and Wales with rapid access to consistent professional health advice and information about health, illness, and the National Health Service (NHS).

It was also hoped that NHS Direct would reduce demand on other primary and secondary care NHS services. NHS Direct was established in 1998 as a pilot project covering 3 areas of England but expanded to cover all of England and Wales by 2001. The service now operates 365 days per year, 24 hours per day, from a network of 21 sites in England with a single site covering all of Wales, and is the world's largest online provider of health care advice answering approximately eight million calls per year. NHS Direct also delivers a website (NHS Direct Online receiving over one million visits per month - http://www.nhsdirect.nhs.uk/), an interactive digital TV channel, and a free health self-help guide.

NHS Direct nurses use clinical decision support software (the NHS Clinical Assessment System (NHS CAS)) to triage (not diagnose) callers' health problems and queries. The NHS CAS is based around approximately 230 computerised clinical algorithms (tree-like structure of mainly yes/no answers). Nurses use the most appropriate algorithm and their clinical judgement to triage each call. The range and clinical severity of the callers reported symptoms dictate the algorithm used (e.g. vomiting, heat stroke, back pain) and the call outcome (e.g. General Practitioner or Accident and Emergency department referral, or self-care advice). Data about NHS Direct calls were originally stored in separate databases residing within the network of NHS Direct call centres, but are now kept in a single national database: the NHS CAS data warehouse. This data warehouse is unique in the UK as it contains timely data about a spectrum of health problems reported throughout England and Wales. In addition to the core clinical algorithms used by NHS Direct there are a number of emergency pathways that remain dormant in the system until activated in emergencies. These pathways will be used by NHS Direct nurses for handling public concern around known releases of chemical/biological agents or incidents involving SARS or pandemic flu. Emergency pathways were designed to distinguish between high and low risk exposure, provide advice and reassurance, and to direct individuals to appropriate centres for assessment and vaccination. Data about emergency pathway usage has not been used for surveillance purposes as, at the time of writing, they have not been used operationally.

Who uses NHS Direct and for what reason?

To assess whether telephone triage data are potentially useful for surveillance purposes, it is necessary to examine to what extent NHS Direct callers are representative of the general population. Under or over representation by particular age groups, geographical areas, or ethnic groups may introduce bias when examining the epidemiology of infectious disease syndromes reported to NHS Direct.

It is estimated that one quarter of the population has used NHS Direct (Knowles et al., 2006). The eight calls per annum are answered by NHS Direct call handlers or trained nurses, compares to eight million out-of-hours calls to GPs prior to the introduction of NHS Direct (Salisbury et al., 2001), 16 million visits a year to Accident and Emergency Departments (DH, 2008), and 190 million consultations with General Practitioners (GPs) (Rowlands and Moser, 2002).

Main reasons for calling NHS Direct

Seventy percent of NHS Direct callers request assessment of symptoms (symptomatic calls), with another 15% receiving in-depth health information and the remaining 15% receiving simple health information (e.g. contact information for local services) or deciding during the call that they do not require the service. Table 1 shows the 10 most used NHS CAS clinical algorithms during March 2006, which accounted for 29% of total symptomatic calls. This list highlights the prevalence of gastrointestinal and respiratory illness reported to NHS Direct (some of which could be indicative of infection) and the high usage of algorithms designed specifically to triage illness in young children. A study of the impact of infections on primary care found that nearly a third of GP consultations were caused by infections, with the highest rates in children <16 years (Fleming et al., 2001). Of the NHS Direct symptomatic calls, approximately 20% receive advice on self care or a pharmacy referral, 50% are advised to see their GP, 8% are advised to visit an accident and emergency department, and for 5% of calls a 999 call (paramedic call out) is advised or made by NHS Direct. The remaining 17% of callers are directed to a variety of other services (e.g. poisons centre, family planning clinic).

Temporal variation in calls

Similar to Accident and Emergency department admissions there is a distinct hourly peak in weekday calls to NHS Direct between 6 and 8 P.M (Munro et al., 2001). This reflects the out-of-hours nature of NHS Direct as call rates also rise at weekends (peaking 9-12AM and 5-8PM), and during national public holidays when doctors' surgeries are generally closed. Call rates are higher during winter, caused by the annual seasonal rise in respiratory disease and common gastroenteric viruses.

Age and sex of NHS Direct callers

The highest call rates to NHS Direct are about young children (<1 year: 358 calls per 1,000 per year; 1-4 years: 173 per 1000 during 2005) and young adults: 76 per 1000). Call rates fall thereafter with increasing age with the lowest call rates from old people. This pattern is consistent with telephone triage data from Australia (Turner et al., 2002). Women are more likely than men to use the service: the ratio of female to male calls (all ages) is 1.3:1. The distribution of calls by age and sex is comparable to that of use of GP services for those under 65 years (Munro et al., 2001). Call rates drop rapidly above 65 years, in contrast to GP services where demand is high in this same age group. The low call rate from the elderly is thought to reflect the older generation's reluctance to use the telephone as a means of accessing out-of-hours care (Foster et al, 2001) or seeking help. The main reason given by individuals as to why they do *not* use NHS Direct is that they 'prefer to see doctors face to face' (Awareness tracking, 2003).

Social deprivation and ethnicity

The small amount of published work about the effect of social deprivation on NHS Direct usage is inconclusive. Questionnaire studies have indicated low usage of the service among lower socioeconomic groups (Knowles et al., 2006; Ring and Jones, 2004), whereas ecological studies show high call rates from areas where these same socio-demographic groups live (Burt et al., 2003, Cooper et al., 2005). An explanation for these seemingly contradictory findings could be that NHS Direct call rates are higher in deprived areas due to higher area level disease but that within any area the affluent individuals are more likely to use NHS Direct. With respect to ethnicity, a report from the UK National Audit Office, published in 2002, suggested that NHS Direct was used less by ethnic minorities (NAO, 2002). Recent NHS Direct figures, however, indicate that the proportion of callers from different ethnic groups reflects the underlying census population. One exception is an under representation from the Chinese sub-group (Chinemana, 2004) who as a group also use traditional Chinese medicine. To limit language barriers to accessing the service NHS Direct has an interpreting service where by callers may state their preferred language and wait for an interpreter who will provide the requested health information.

Other factors that may affect demand

It is possible that the perceived usefulness and safety of NHS Direct may influence demand for the service. Evidence suggests, however, a high service satisfaction rate: 97% in a questionnaire survey of over 1000 callers (Munro et al., 1998)), and a self reported compliance rate with advice given of 75% (Munro et al., 2001) and 64% (Foster et al., 2003) in separate studies. With regards to safety, when researchers logged serious adverse clinical outcomes of NHS Direct calls, as reported by coroners, A&E consultants, press reports and NHS Direct centre managers, only 3 cases out of 280,000 calls were found. Three equates to a 'critical event rate' of 0.001% (Munro et al., 2001). Foster found that a small proportion of callers (0.3%, 15 callers) who were advised not to attend A&E were subsequently admitted although larger studies have shown nurse consultation is not associated with an increase in adverse events (Lattimer et al., 1998; Lambell et al., 2003). It should be remembered that NHS Direct nurses triage more than half of callers for treatment elsewhere thus reducing the likelihood of serious illness going untreated. The use of nationally agreed and written protocols, rather than unstructured triage, may also improve safety (Cartmill et al., 2001) and reduce the likelihood of the 'missed case'.

Table 1. Ten most used CAS clinical algorithms (March 2006) by the 21 NHS Direct sites in England

10 most used clinical algorithms	% total clinical algorithms used
Abdominal pain	5.8%
Vomiting	3.0%
Toothache	2.9%
Vomiting, toddler (Age 1-4 years)	2.7%
Fever, toddler (Age 1-4 years)	2.6%
Chest pain	2.5%
Diarrhoea	2.5%
Headache	2.5%
Diarrhoea, infants and toddlers (age 0-4 years)	2.4%
Sore throat	2.4%
Total	29.3%

Source: NHS Direct (April 2006)

All NHS Direct call centres use the same computerised clinical algorithms so health advice and information should be provided in a consistent way across the country. However, there may be variation in the way in which the system is used between individual nurses as it has been demonstrated that nurses use rule based systems in a variety of ways, using their own experience to deliver a personalised service (Greatbatch et al., 2005). Indeed, the clinical experience of NHS Direct nurses affects the outcomes of calls (Lambell), with experienced nurses more likely to provide self-care advice (O'Cathian et al., 2004) than less experienced colleagues. Although this factor may explain differences in call outcomes between individual nurses there is no published evidence to suggest that it causes variation in call classification and outcomes between NHS Direct call centres. Inter-nurse variation should therefore not cause geographical variation in call rates for individual syndromes.

It is important to also consider whether the quality and availability of existing or alternative health services influences the geographical variation in NHS Direct call rates. Unfortunately, there is no published research addressing this question, despite significant local variation in call rates (Cooper et al., 2005; Burt et al., 2003) and the type and structure of GP out-of-hours care (Pooley et al., 2003). Turning this question around, research studies have demonstrated that the introduction of NHS Direct has been associated with a small (8%) reduction in calls to other out-of-hours primary care services (GP cooperatives) but has had no effect on the numbers of ambulance calls and emergency department attendances (Munro et al., 2005). NHS Direct thus operates more as an additional tier of health care than as a replacement of existing health care services. Finally, a questionnaire survey found that those with restricted access or difficulties in using a telephone were less likely to use NHS Direct (Knowles et al., 2006). A dedicated text phone service for the deaf and hard of hearing does exist for this group.

Disease surveillance

Definition

The current concept of disease surveillance, as a means of monitoring population rather than individual health, was proposed by Alexander Langmuir of the Centres for Disease Control in the 1950s, defined as follows (Langmuir, 1963):

> *Surveillance, when applied to a disease, means the continued watchfulness over the distribution and trends in the incidence of disease through the systematic collection, consolidation and evaluation of morbidity and mortality reports and other relevant data.*

Put more simply, disease surveillance provides information for public health action to protect and promote human health. This action should lead to the alerting of health professionals, the government, and the public to new disease threats. For example, surveillance may identify a localised or national increase in infectious disease that in turn leads to contact tracing of those possibly affected by the outbreak and a targeted intervention to protect these individuals and prevent further spread of disease (e.g. legionnaires' disease (Ruscoe et al., 2006), salmonella infection (Gillespie and Elson, 2005)). Alternatively, action resulting from surveillance may occur over a longer time period, for example the implementation of a national vaccination programme to counter an increase or re-emergence of a vaccine preventable disease (e.g. pertussis (Miller and Farrington, 1998) or meningococcal C (Shigematsu et al., 2001)). Other aims of surveillance include the clarification of clinical features (e.g. H5N1 influenza virus (Tran et al., 2004)), risk factors (HIV (Bodley-Tickell et al., 2004)) and outcomes. This raises hypotheses for further investigation and supports ongoing tracking of health outcomes to assist evaluation of policies and interventions.

A brief history of surveillance

> *For thousands of years, epidemics were looked upon mainly as divine judgments on the wickedness of mankind, and it was believed that the punishments were to be avoided by appeasing the gods (Rosen, 1958).*

In the modern era more worldly measures, including surveillance, are also used to prevent and control epidemics. Mortality and morbidity records, however, have been used to inform public health action in Europe since late medieval times, well before the modern concept of disease surveillance existed. Cause specific mortality records were used in London from the 17th century to issue John Gaunt's 'Bill of Mortality' (an analysis of mortality that used novel statistical methods to develop innovative concepts such as life expectancy tables). During the 18th Century the idea that surveillance could be used as a tool for providing information on population health gained prominence in many European countries. For example, in 1766 Johan Peter Frank developed a system of 'medical police' in Germany which incorporated information on school, maternal and child health, injury prevention, sewage and water treatment. Central to this movement was the concept that the effects of disease were important to the individual *and* to the community. In North America, by the 1740s, Rhode Island tavern keepers were required to report to the authorities' cases of small pox, cholera and yellow fever among their patrons. Accurate and complete mortality data for the UK was collected by the General Registrar Office from 1836 from which William Farr applied early medical statistics to develop what might be considered a modern surveillance system. During the 20th century many different surveillance systems have developed encompassing a variety of data sources, collection methods, analytical techniques and means of dissemination.

Surveillance in the present time

Today the World Health Organisation is responsible for collecting infectious disease data and administering the International Health Regulations (IHR) which, through international law, define a global public health protection strategy. This has recently taken a major step forward with the new International Health Regulations (IHR)

(Baker and Fidler, 2006) which came into force in the summer of 2007. The IHR are an improvement on the previous system in that they are established to detect all international health threats (including new ones like SARS), and are not just a limited list of specific infections. Also they require that countries establish adequate surveillance systems to perform this role. A central pillar is the requirement of a global surveillance system with the ability to detect and report public health emergencies of international concern. With this in mind, the WHO and the European Centre for Disease Prevention and Control (ECDC) are collaborating with the aim of developing a public health alerting system for the European Union (Kaiser and Coulombier, 2006). For this a strong emphasis is placed on epidemiological surveillance, early warning and response. This will rely heavily on the strengths of individual EU member states surveillance networks and must also take advantage of European wide surveillance networks coordinated from within single member states. For example, the European Influenza Surveillance Scheme (EISS, 2008), which will become part of ECDC, collects and exchanges timely clinical and laboratory information about influenza from 26 European countries (Meijer et al., 2006). Enter-net (already integrated into ECDC), an international network including 15 EU countries, provides surveillance of human gastrointestinal infections (Enter-net, 2005). A coordinated approach to surveillance may be difficult to achieve, however, given large differences between 'epidemic intelligence' gathering within European countries.

The Health Protection Agency in the UK

Historically, the NHS provided health protection services in the UK through local Health Authorities (e.g. immunisation programmes, infection control teams and local Consultants in Communicable Disease Control (CCDCs)). However, recent events of national and global importance (2001 UK foot and mouth outbreak and the 2001 terrorist attacks on the US) convinced the UK Chief Medical Officer (Sir Liam Donaldson) that an integrated approach to health protection and emergency planning was required. Thus the Health Protection Agency for England was established in 2003 following the publication of 'Getting Ahead of the Curve' (DH, 2002). The role of the HPA is to protect the health of the community against biological, chemical, radiological and environmental hazards. It also advises the Government on emergency planning and health protection policies, provides advice and information to health professionals and the public, and responds to emerging public health threats. Health protection surveillance of infectious disease and environmental hazards is a core part of the work of the Health Protection Agency.

The HPA is organised through national, regional and local tiers. On a national level the work of the HPA is carried out by the Centre for Infections, Centre for Radiation, Chemical and Environmental Hazards, and the Centre for Emergency Preparedness and Response. HPA Regional services (through 9 regional offices) provide specialist epidemiology (and surveillance), microbiology and emergency planning services, help manage major incidents, and coordinate the activities of local Health Protection Units (HPUs). The HPUs are the local tier and 'front line' of the HPA and work directly with their local PCTs (and hospital trusts) to provide a full range of health protection services including health emergency planning, immunisation, infection control and HIV prevention. Other agencies also have a key role in health protection liaising with the HPA in a number of ways. For example, environmental health officers, based within local Government, investigate food or water borne infections, and enforce food safety legislation. The Food Standards Agency warns the public of food hazards and provides a Government response to major food-borne outbreaks.

Health protection surveillance in the UK

Health Protection surveillance is the surveillance of infectious diseases and environmental hazards that may pose a threat to public health. Its purpose is to:

- identify individual cases of disease or outbreaks so action can be taken to investigate and prevent spread or further cases,
- measure the incidence of disease to detect change and implement control measures,
- evaluate these control measures, and
- detect and describe new infections or environmental hazards so as to identify risk factors and aetiology.

In the UK a variety of data sources are used by the HPA for infectious disease surveillance, including hospital and primary care clinical information systems, laboratories, schools, the Office for National Statistics, and Environmental Health Departments, described below:

- *NOIDS*: Any clinician suspecting a diagnosis of a statutory notifiable infectious disease (NOIDS) (e.g. anthrax, cholera, measles) is legally required to report it to the local authority who in turn report to the Health Protection Agency. NOIDS data provide an effective way of monitoring diseases where laboratory tests are not necessarily performed (e.g. pertussis).
- *Laboratory data*: Laboratory testing of a limited number of infections is reported through NHS and private laboratories and is published in a national weekly Health Protection Report (HPA, 2008).
- *Hospital data* provide information on severe disease that requires hospitalisation. These data are often not available with diagnoses in a timely enough fashion for routine surveillance systems.
- *Death certifications* can be useful for the surveillance of infectious diseases which may lead to or cause death, and be coded on death certifications (e.g. meningococcal disease).
- *Enhanced surveillance* systems exist for infections or hazards of particular public health importance such as meningococcal disease, antimicrobial resistance, or travel associated legionella infections. Laboratory data for these diseases may be combined with further epidemiological data collected from other sources (e.g. clinicians, patients).
- *Sexually transmitted infections* data reported to Genitourinary Medicine clinics are collected and usually released in aggregated format (K60 returns).
- *The Medical Officers of Schools Association* (MOSA) report weekly illness from 55 boarding schools in England and Wales. These data are particularly useful for early identification of influenza-like-illness outbreaks.
- *Primary care data (discussed below).*

The surveillance of infectious diseases is not consistent in terms of the levels of scrutiny placed on each disease. Some common infections which cause substantial morbidity and economic cost are rarely tested for in laboratories. For example, information about an estimated 1 in 1600 cases of norovirus (a common cause of viral gastroenteritis) reach HPA surveillance systems, whereas far more complete surveillance data exist for other infections (e.g. salmonella: 1 in 3.2 cases) (Wheeler et al., 1999). Other infections have a number of data sources from which to produce surveillance reports. For example, because of the unpredictability and ever changing nature of influenza (due to antigenic drift and shift of viruses), the influenza surveillance programme uses data from laboratories, death certificates, MOSA, General Practitioners and NHS Direct, while HIV and STI surveillance is based on an even greater range of sources (HPA, 2008).

Chemical incidents are unforeseen events leading to population exposure and/or illness relating to non-radioactive or toxic substances. The number of individuals who suffer health effects from incidents typically varies from 2-50; the most commonly reported causative chemicals are asbestos and carbon monoxide (Hawker et al., 2001). The HPA has responsibilities for the health aspects of chemical incidents monitoring them from a range of sources, including local Health Protection Units, the Fire Service, and the National Poisons Information Service (NPIS). The NPIS conducts surveillance on the toxic effects of products in the United Kingdom, and can provide enhanced surveillance about enquires regarding specific products (e.g. aerosol sprays (Tizzard et al., 2003). There are currently no timely local surveillance data monitoring the common health effects of chemical incidents (e.g. respiratory problems, eye irritation) in the UK.

Primary care surveillance in England and Wales

It is estimated that 40% of the UK population consult their general practitioner (GP) at least once during the year specifically because of infection (Fleming et al., 2001). With this in mind, since 1998 the HPA has taken an active role in primary care epidemiology and surveillance (McNulty and Smith, 2000). The HPA West Midlands Regional Surveillance Unit is home to the Primary Care Surveillance Team which takes a national lead for the HPA Centre for Infections on health protection surveillance in primary care. The surveillance team use symptoms reported to NHS Direct and diagnoses made by GPs for surveillance purposes. The three national sentinel GP surveillance systems (there is a separate GP surveillance scheme in Wales) in England and Wales are:

1. The Royal College of General Practitioners Weekly Returns Service (RCGP): A clinical information system that collects twice weekly data (region, age and gender specific) on every new episode of disease diagnosed in a network of 92 General Practices covering over 800,000 people located throughout England and Wales. The unit is unique in the UK in that it has collected data since the 1970s and can monitor acute and long-term trends in a wide range of infections and medical conditions (Fleming, 1999). A sub-set of RCGP practices also collect nose and throat swabs from patients for virus isolation and characterisation of influenza and respiratory syncytial virus (RSV).

2. Q-Research: A database of the health needs, risks, care and outcomes of over 7 million patients registered with approximately 500 practices across the UK. As well as diagnoses this system is also able to collect and present timely information on diagnoses linked to prescriptions and outcomes (e.g. influenza with prescribed anti-virals, impetigo with fusidic acid, emergency hospital admissions for asthma) (Harcourt et al., 2005).

3. QFLU: A daily data collection and analysis service containing data from 2700 general practices in the UK, covering a population of almost 17 million patients. Data are available on consultations for influenza and pneumonia, antiviral prescriptions, and influenza vaccine uptake (Hippisley-Cox et al., 2006). Both Q-Research and Q-Flu use data from GP surgeries using the EMIS computer system to hold patient records and prescriptions.

The benefit of primary care surveillance systems is that diagnoses of non-notifiable diseases, for which there may be no laboratory confirmation, can be related to a defined population (calculated from GP lists held within computerised patient administration systems). Along with data from NHS Direct (discussed later), the GP surveillance systems described above complement each other to provide a timely national picture of disease in primary care in England and Wales. Table 2 summarises the properties of primary care surveillance systems from which the comparative advantage of these systems becomes apparent. For example, the RCGP has a 30 year baseline from which to interpret current trends, and has an integrated microbiological testing scheme. The QResearch/QFlu databases link diagnoses to prescriptions, vaccine uptake and social factors, so are of use for surveillance *and* research purposes.

Table 2. Attributes of primary care surveillance systems used by the HPA: the NHS Direct syndromic surveillance system, the Royal College of General Practitioners Weekly Returns System (RCGP) and the QRESEARCH/QFLU surveillance systems.

Attribute	NHS Direct	RCGP	QRESEARCH/QFLU
Population coverage	England and Wales (52 million)	90 practices (860,000)	3,039 practices (20 million)
Timeliness	+++ (daily)	++ (twice weekly)	+ (weekly)
Smallest geographical unit data routinely available to	+++ (postcode district)	+ (Region)	++ (Primary Care Trust)
Historical baselines	++ (6 years)	+++ (>30 years)	++ (5 Years)
Alerts of common conditions	++	+++	+++
Alerts for deliberate release (prodromal)	+++	++	+++
Diagnoses	-	+++	+++
Integrated microbiological testing	+	++	-
Linked prescribing data	-	-	++
Vaccine uptake data	-	-	++
Risk group analyses (e.g. deprivation, age-group)	++	-	+++

Key: - N/A + poor utility ++ reasonable utility +++ good utility

+++ holds comparative advantage over other systems

The range of surveillance systems used by the HPA must evolve if it is to mount an effective response to emerging or new infections and political needs brought about by the threat of bio-terrorism. Currently it may take weeks for a complete record of laboratory samples to reach surveillance systems. This reduces the likelihood of rapid detection and response to an emerging public health problem. There is also bias inherent within laboratory surveillance systems both in terms of the type of infections that are likely to be tested for, and the demographic group that provide the samples (e.g. hospitals are more likely to refer samples from children than adults for laboratory tests). With regard GP surveillance systems, data are generally reported on a weekly basis and by geographical units of many millions of people. A rapidly evolving local infectious disease outbreak may therefore go undetected. Once a local problem is detected (possibly a few weeks later through laboratory surveillance) GP surveillance systems have insufficient geographical granularity to accurately characterise the event as it unfolds. Although the main sources of UK infectious disease surveillance data are likely to remain laboratory and clinician based, new sources such as NHS Direct should be explored as a potentially useful adjunct.

NHS Direct calls may in particular have added value because:

- **NHS Direct is available to the entire population of England and Wales.**
- **The data are available on a daily basis, and are thus making it suitable for real-time surveillance.**
- **The data cover a wide range of syndromic conditions, not all monitored elsewhere.**
- **NHS Direct is many peoples first or only point of contact with the health service, providing an early opportunity to identify an increase in illness.**
- **The system is centrally managed, which allows for central quality control and rapid modifications when needed.**

Syndromic surveillance

Rationale

Syndromic surveillance is the collection and analysis of non-specific or pre-diagnostic data for monitoring population health and informing public health action. A widely used definition reads:

> *An investigational approach where health department staff, assisted by automated data acquisition and generation of statistical alarms, monitor disease indicators in real-time or near real-time to detect outbreaks of diseases earlier than would otherwise be possible with traditional public health methods (Buehler et al., 2004).*

Syndromic surveillance systems analyse daily or real-time data with the aim of providing early warning of potential health protection issues (e.g. infectious disease outbreaks and epidemics, bio-terrorist attacks, or major chemical incidents). It is hoped that this will alert public health officials of problems before they would have been alerted via traditional surveillance sources. Figure 1 illustrates how, after a deliberate release of a bio-terrorist agent, there would be a rapid increase in those presenting with symptoms. This may be a single symptom or more complex syndrome. It is the job of the syndromic surveillance system to identify this increase before a rise is detected in those presenting with the later and more serious stages of disease. This lead-time (t in Figure 1) enables a more rapid public health response, thus reducing morbidity and mortality. Additional benefits of syndromic surveillance are the timely acquisition of data for monitoring known outbreaks, the strengthening of public health networks as a result of the follow-up of syndromic surveillance ‘signals’, the identification of secular patterns of disease, the flexibility to be ‘tweaked’ to detect new threats (e.g. heat waves, or poisonings from new market products), and the provision of reassurance when significant trends are not detected during times of perceived high risk.

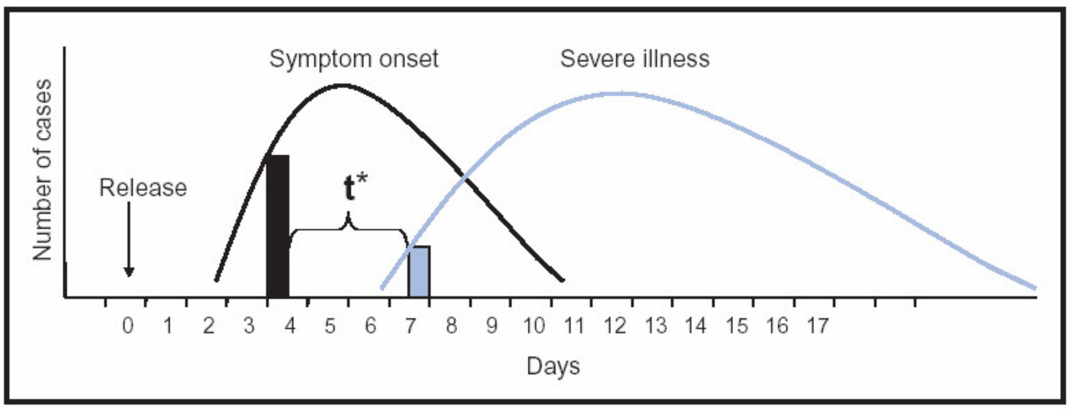
Syndromic surveillance systems are distinguished from other surveillance systems by the types of data (syndromic rather than diagnostic) that are monitored and the timeliness of the analyses and dissemination of results. Another distinction is the range of disciplines from which expertise is drawn to operate such systems, including epidemiology, medicine, statistics, bio-statistics, human geography and informatics. The main types of data used are ambulance encounters (Lazarus et al., 2001) and hospital admissions (Heffernan) (used for approximately 60% of U.S systems (Sosin and DeThomasis, 2004), clinician visits (Miller et al.,2004), and pharmacy sales (Das et al., 2005), although calls to health help-lines (‘nurse hotlines’) (Baker et al., 2003), laboratory test orders (Ma et al., 2005) and work absenteeism reports (Heffernan et al., 2004) have also been used. Prescription data (Chen et al., 2005), poisons control centre data (Derby et al., 2005), medical billing data (Sloane et al., 2006), self reporting via emails (Marquet et al., 2006) and internet forms (Wethington and Bartlett, 2006) have also been proposed as a means of monitoring population health.

Much of the development of syndromic surveillance systems has been in the US (approximately 100 systems in 2003 (Buehler et al., 2003)) with city (Micheal et al., 2002; Lazarus et al., 2002; Heffernan et al., 2004) or county wide (Paladini, 2004) schemes the norm. An ambitious national system – *BioSense* (Broome, 2004, Sokolow et al., 2005) - now also aims to improve US capabilities for real-time syndromic surveillance through access to data from sentinel networks of hospitals, health care facilities and laboratories across the US. Elsewhere syndromic surveillance has been used prospectively in a variety of settings, for example around the recent G8 conferences in Japan (Osaka et al, 2002) and the 2003 Rugby World Cup in Australia (Muscatello et al., 2005), for surveillance of West Nile Virus in the

Netherlands (Rockx et al., 2006), and for a variety of syndromes reported to General Practitioners in New Zealand (Jones and Marshall, 2004).

Figure 1. Rationale for syndromic surveillance

Source: Henning, 2004



* t = time between detection by syndromic (prediagnostic) surveillance and detection by traditional (diagnosis-based) surveillance.

Usefulness

Syndromic surveillance is a relatively young field and was boosted in the United States by the terrorist attacks of September 2001 and by the deliberate domestic releases of weaponised anthrax directly afterwards. These events generated substantial investment in the US to develop systems for the early detection of the health effects of bio-terrorist attacks. Consequently syndromic systems were established in advance of the usual evaluation and small scale pilot studies that precede large scale investment. Since 2001 there have been many published time series and simulation studies suggesting that there are potential benefits of using syndromic surveillance for early warning of bio-terrorist attacks (Nordin et al., 2005), or rises in naturally occurring infections (Lazarus et al., 2002). It is yet to be demonstrated, however, that syndromic surveillance systems can prospectively provide *consistent* early warning of localised outbreaks of disease.

There is much published work about the New York City Syndromic Surveillance system, in operation since 1995. Here the main reported gains are in providing the earliest indication of city-wide influenza activity and also of increases in gastrointestinal disease in advance of a rise in institutional outbreaks of norovirus (Heffernan et al, 2004). This system also provided reassurance during March 2003 that a rise in febrile illness was not caused by SARS (Steiner-Sichel L et al., 2004). Also in New York City, syndromic surveillance identified an increase in diarrhoeal illness prompting an investigation that identified otherwise unreported disease, caused by the consumption of bad meat that had spoiled after a power cut (Marx et al., 2006). In the main, however, syndromic surveillance has not been useful in detecting acute localised outbreaks of GI illness in New York City (Balter et al., 2004). Equally it is difficult to find published examples of events detected via syndromic surveillance (in New York or elsewhere) that have led to public health action. A review focussing on the efficacy of syndromic surveillance systems for waterborne disease detection found no evidence of earlier detection or a reduction in the effect of outbreaks (Berger et al., 2006).

Evaluation

As a result of inconclusive evidence about benefits of syndromic surveillance some have questioned the wisdom of the substantial investment made in recent years (Reingold, 2003). Academics have pointed out the difficulties of measuring the impact of syndromic surveillance systems in a scientific way, for example using sensitivity, specificity, and positive predictive value (Bravata et al, 2004). Other criticisms note that:

- syndromic surveillance can only detect events affecting large numbers of people,
- investigating false alarms is costly,
- in the absence of known bio-terrorist attacks there is no reference standard from which to evaluate sensitivity,
- patients with a particular illness will present with a variety of syndromes/prodromes and be coded in different ways thus diluting any signal, and
- surveillance ‘signals’ rarely initiate further testing of patients for definitive diagnoses.

A review of peer-reviewed articles (to April 2002) about systems designed to detect bio-terrorist events, supports some of these criticisms, highlighting a lack of information about both the sensitivity of syndromic surveillance systems and how they have facilitated decision making (i.e. their usefulness) (Bravata et al., 2004). The authors concluded that there was little scientific evidence to guide those managing these systems. This is born out by a summary of the 99 abstracts submitted to the 2003 national syndromic surveillance in the US (Sosin and DeThomasis, 2004). Only 7 abstracts reported evaluation of investigations resulting from syndromic surveillance 'signals'. The article suggested that a simplified version of the CDC framework - focussing on timeliness, validity, and *usefulness* – would help measure how public health is improved through syndromic surveillance.

Two surveillance systems have published results of using the 'Framework for Evaluating Public Health Surveillance Systems for Early Detection of Outbreaks' (Beuhler et al., 2004). The first used semi-structured interviews with a small group of stakeholders of the NHS Direct syndromic surveillance system. These stakeholders believed that the system was useful for detecting national increases in norovirus and influenza-like-illness, tracking epidemics, reassuring the public, augmenting other surveillance systems, and generating research work (Doroshenko et al., 2005). Limited results from the second study (evaluating Maryland's ESSENCE system) cited usefulness for determining when to initiate and cancel a county influenza vaccination program (Lombardo et al., 2004). A further study using structured discussions with users of ESSENCE highlighted its usefulness for tracing disease patterns over time, providing a timely geographic/demographic picture of community disease incidence, and determining the impact of health interventions (Hurt-Mullen and Coberley, 2005). Details of these reported benefits were absent from the report however.

Due to the large amounts of daily syndromic data analysed, syndromic surveillance systems inevitably generate many statistical anomalies ('signals'). The focus on analysis of daily data to a local level also means that signals are often based on very small numbers of cases. A high proportion of which, on investigation, turn out to be false positives (Baker et al., 2004, Heffernan et al., 2004). A clear process is therefore needed to decide which signals to investigate further and which to discard. Issues such as the quality and consistency of data, and the background and seasonal variation in disease must be considered before decisions are made to investigate further (Figure 2). These decisions are crucial to the ultimate success of syndromic surveillance systems.

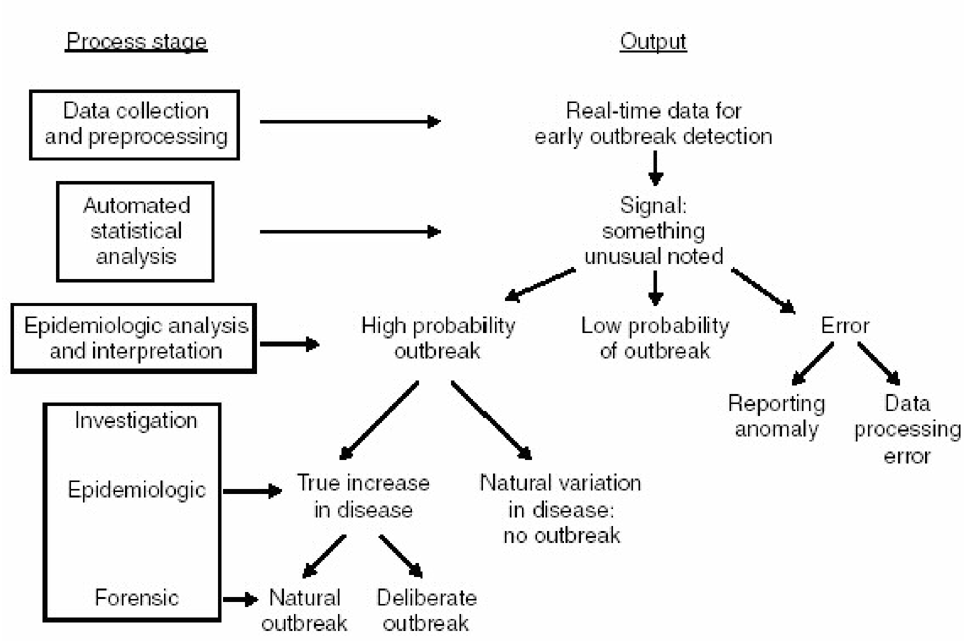
As well as epidemiological issues, those handling syndromic data must also consider patient confidentiality, data protection law, data ownership, and administrative accountability (i.e. local or national responsibility). It is also important to understand and distinguish between expected data anomalies (e.g. caused by influenza-like-illness) that influence well-established public health responses, and more unusual anomalies (e.g. un-seasonal or caused by severe clinical outcomes) which may trigger unfamiliar response mechanisms. In short, a challenge for those working in the field of syndromic surveillance is to define the types of public health interventions that may occur following syndromic surveillance signals, and the likely outcomes of these responses and interventions (Sosin, 2003). This will help clarify where the added value of syndromic surveillance lies.

Previous work using telephone triage data for syndromic surveillance

In 1998 a retrospective time-series analysis revealed a 17-fold increase in nurse hotline calls about diarrhoea during a citywide outbreak of cryptosporidiosis in Milwaukee (Rodman et al., 1998). Two retrospective analyses of telephone triage data from the Baltimore - Washington metropolitan area demonstrated that individual calls can be used to predict respiratory and gastrointestinal final diagnoses (Henry et al., 2006), and that trends in telephone triage data accurately predict trends in doctor diagnoses (Magruder et al., 2005). This is supported by a study that concluded that rises in respiratory and gastrointestinal calls to the Scottish national telephone triage system heralded later rises in influenza-like-illness and norovirus reports respectively (Wilson et al., 2006). Studies comparing US telephone triage data against CDC influenza surveillance data produced conflicting results, one study showing a peak in local primary care centre telephone triage data after laboratory and clinical data (Espino et al., 2003), the other showing a peak in triage data at the same time (Yih et al., 2006). Neither of these studies looked at the early warning potential of these data in terms of alerting when baseline activity was exceeded and significant influenza activity was likely to follow. In Canada, an evaluation of data from the Ontario's Telehealth helpline (Rolland et al., 2006) shows some promise for surveillance of respiratory disease (Rolland, 2007) although prospective analyses is not yet operational. The paucity of published work about the use of telephone triage data for syndromic surveillance provides an opportunity, within this thesis, for an original evaluation using NHS Direct data.

Indeed, there has been a call to explore further the use of novel types of data for surveillance and to establish what sort of health behaviours (e.g. making a telephone call) serve as early indicators of community morbidity (Sosin, 2003).

Figure 2. Process model for early outbreak detection using syndromic surveillance.



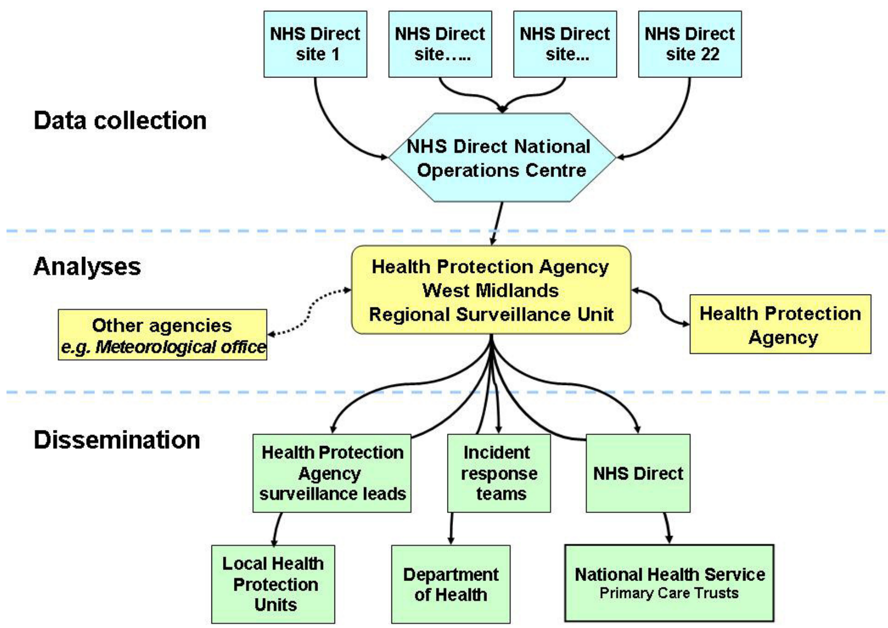
Source: Buehler et al., 2002

Why use NHS Direct data for syndromic surveillance?

NHS Direct sites use clinical software (NHS CAS) and reporting tools capable of recording and reporting caller details on a daily basis. This timeliness, coupled with the national coverage of NHS Direct, the central standardisation and the wide range of recorded syndromes, make the data suitable for population health surveillance. Where a deliberate release of a harmful agent causes illnesses with a mild prodromal phase, a proportion of people may contact NHS Direct before visiting other health services. Consequently there may be an opportunity to identify an increase in illness before it is reported to other primary care or secondary care services. The rationale behind using telephone triage data from NHS Direct is that this information may provide the first signs of an emerging public health problem.

The aim of the NHS Direct syndromic surveillance system is to identify an increase in syndromes indicative of common infections and diseases, or the early stages of illness caused by the deliberate or accidental release of a biological or chemical agent.

Figure 3: National NHS Direct syndromic surveillance system flowchart.



The monitoring of syndromes reported to NHS Direct could help track a rise in illness over time, by characterizing the age, geographical location and syndromes of affected callers. Since 2001 the national NHS Direct syndromic surveillance system has collected daily telephone triage data from 22 NHS Direct sites for prospective daily surveillance of 11 syndromes. These data are analysed by the HPA West Midlands Regional Surveillance Unit (performing a national primary care surveillance role) each weekday and interpreted with assistance from other departments of the HPA and other agencies (Figure 3). Rises in syndromes considered of public health importance are reported to local or national health protection teams. A weekly surveillance bulletin is distributed across the HPA and within NHS Direct, is widely distributed across the NHS and Department of Health, and is published on the internet (http://www.hpa.org.uk/infections/topics_az/primary_care_surveillance/NHSD.htm).

A small specialist surveillance team manage surveillance activities, with a multi-disciplinary collaborative group (HPA and NHS Direct staff) providing strategic direction, additional technical and scientific support, and co-ordination around other non-surveillance projects.

The surveillance system is now well established, with over seven years of operation. This provides a sufficiently long enough period from which to critically examine a number of published NHS Direct syndromic surveillance reports. Given that there are at least 100 disease surveillance systems alone in England and Wales, this thesis also provides an opportunity to discuss the added value of using NHS Direct data for health protection surveillance in the UK. The likely continuation of the system means that recommendations can be made for developing the system in the UK, and for an international audience potentially considering using telephone triage data for surveillance purposes.

Research question

This primary aim of this thesis is to describe and evaluate the use of NHS Direct call data for surveillance purposes. The principle research question is: "What is contribution of NHS Direct data to health protection surveillance in England and Wales?" The research questions for each study (Chapters 2-11) are given in Table 3.

To explore these research questions NHS Direct syndromic surveillance data are analysed with reference to the commonly held principles and objectives of disease surveillance systems held by the Health Protection Agency of England and Wales (HPA, 2005).

The principles are that surveillance systems have the following characteristics:

- Exploit existing data already collected for other purposes.
- Have clearly defined aims.
- Be representative of the population under surveillance.

- Involve data providers in the operation of the system and ensure that they receive surveillance outputs.
- Share data within their own and other organisations.
- Have accessible and clearly defined system outputs.
- Lead to action in order to achieve public health goals.
- Achieve acceptable sensitivity, specificity and positive predictive value.
- Conduct regular evaluation to ensure that the system remains fit for purpose.
- Be flexible enough to respond to new public health threats.

The objectives are that surveillance systems:

- Detect outbreaks of respiratory or gastrointestinal disease incidence.
- Record significant trends in disease.
- Provide early warning of a rise in illness in the community.
- Validate and support other surveillance systems.
- Anticipate disease threats through modelling (e.g. for pandemic flu).
- Ascertain cases of public health importance.
- Estimate the burden of disease in specific populations.
- Raise hypothesis for investigation, and facilitate research.

Outline of the thesis

The outline of the thesis is presented in Table 3. After the introductory chapter the body of the thesis is split into three parts. Part 1 (chapters 2-4) introduces the NHS Direct service, describes the nature of NHS Direct data, and the aims and methodology of the NHS Direct syndromic surveillance system. Part 2 (chapters 5-8) contains analytical studies that compare NHS Direct data against data from other surveillance systems. These studies examine the utility of syndromic surveillance for early outbreak detection at a local, regional and national level. Part 3 (chapters 9-11) presents operational results from the syndromic surveillance system and a feasibility study to link syndromic surveillance directly to microbiological testing.

Part 1: NHS Direct data and syndromic surveillance

Chapter 2 briefly describes the NHS Direct service and discusses the nature of NHS Direct data with respect to epidemiological considerations such as data availability and data quality. This chapter also considers factors that will determine to what extent the public health community will use NHS Direct call data for public health purposes.

The impact of age, sex and deprivation on call rates is explored in *Chapter 3*. The importance of this chapter is in determining the representativeness of the surveillance data and whether any potential biases are inherent within the surveillance system.

Chapter 4 introduces the rationale and aims of the NHS Direct syndromic surveillance system. It goes on to outline the statistical methodology used to analyse daily time series data and the public health protocols used to investigate significant rises in syndromes. The outcomes of these investigations are described.

Part 2: Evaluation studies - external validity and comparison against other data sources

Chapter 5 is a simulation study that tests whether a historical cryptosporidiosis outbreak would have been detected by the surveillance system, and whether it would have been detected in advance of the existing alerting mechanism. The study models a range of reporting rates of diarrhoea to NHS Direct so that the likelihood of detecting the outbreak can be studied under different scenarios.

Chapter 6 describes an evaluation study which uses a Poisson regression model of NHS Direct data and GP surveillance data to calculate NHS Direct influenza incidence thresholds that signal the start of the influenza season. It then determines whether, retrospectively and prospectively, these thresholds provide early warning of national influenza outbreaks during five separate winters.

Chapter 7 uses a multiple regression model to estimate the contribution of specific respiratory pathogens to the seasonal variation of NHS Direct calls. It also provides disease estimates for these pathogens as reflected within NHS Direct calls.

Chapter 8 aims to use spatio-temporal cluster detection methods to identify the origins and diffusion of national rises in common viral diseases (influenza and norovirus are used as case studies).

Part 3*: *Surveillance results focussing on respiratory and gastrointestinal infections

Chapter 9 outlines the age distribution and outcome of NHS Direct calls suggestive of gastrointestinal infection. These calls account for one in ten of all calls and one in six of calls about children less than 5 years.

Chapter 10 reports the second year's experience of using NHS Direct calls for influenza surveillance. Daily and weekly trends in age-group specific NHS Direct 'cold/flu' calls are presented. These data are compared against existing sources of clinical and laboratory influenza data to assess how well they describe the influenza season and support other systems.

Chapter 11 discusses the ability of the system to link to laboratory surveillance and identify individual cases of potential public health importance. A pilot study is described that tested the feasibility of NHS Direct callers self-sampling to provide viable samples for influenza culture.

General discussion

The final chapter presents the research findings and discusses the contribution of NHS Direct call data to health protection surveillance in England and Wales. The results of two formal stakeholder evaluations of the system are summarised to support these main findings. Finally, recommendations are made for the UK Health Protection Agency, NHS Direct, research community, and those considering using telehealth data for surveillance purposes elsewhere.

Table 3. Outline of the thesis

	Chapter title	Research question	Design
1.	General Introduction.		
Part 1: NHS Direct data and syndromic surveillance			
2.	NHS Direct derived data: an exciting new opportunity or an epidemiological headache?	What issues are relevant to the use of the data for surveillance and epidemiological purposes?	Discussion paper.
3.	The effect of deprivation, age and gender on NHS Direct call rates.	What is the impact of deprivation, age and sex on call rates at two NHS Direct sites?	Ecological study.
4.	National Symptom Surveillance using calls to a telephone health advice service — United Kingdom, December 2001– February 2003.	How can we use NHS Direct data for syndromic surveillance?	Descriptive study.
Part 2: Evaluation studies - external validity and comparison against other data sources			
5.	Can syndromic surveillance data detect local outbreaks of communicable disease? A model using a historical cryptosporidiosis outbreak.	Would a historical cryptosporidiosis outbreak have been detected using the current NHS Direct surveillance system?	Analytical study (statistical simulation).
6.	Can syndromic thresholds provide early warning of national influenza outbreaks?	Can we define national threshold levels of NHS Direct influenza-like-illness?	Analytical study (*Poisson* regression).
7.	The contribution of infectious pathogens to the seasonal distribution of NHS Direct calls.	What is the contribution of respiratory pathogens to the seasonal variation in NHS Direct calls?	Analytical study (multiple regression).
8.	Tracking the spatial diffusion of influenza and norovirus using telehealth data: a spatio-temporal analysis of syndromic data.	Can spatial analyses of NHS Direct vomiting and fever calls to a local level provide added value for norovirus and influenza surveillance?	Analytical study (spatio-temporal scan statistic).
Part 3: Surveillance results focussing on respiratory and gastrointestinal infections			
9.	What can analysis of calls to NHS Direct tell us about the epidemiology of gastrointestinal infections in the community?	What is the epidemiology of NHS Direct calls suggestive of gastrointestinal illness?	Descriptive analyses.
10.	Use of NHS Direct calls for the surveillance of influenza - a second years' experience.	What is the usefulness of NHS Direct data for influenza surveillance?	Descriptive analyses.
11.	Linking syndromic surveillance with virological sampling: telehealth callers self-sample for influenza.	Can self-sampling by NHS Direct callers provide early warning of an increase in influenza and added value over existing surveillance?	Validity study.
12.	General discussion.		

References

Awareness Tracking. Taylor Nelson Sofres 2003 (internal Department of Health Report).

Baker M, Smith GE, Cooper DL, et al. Early Warning and NHS Direct: a role in community surveillance? J Publ Hlth Med 2004; 25:362-368.

Baker MG, Fidler DP. Global public health surveillance under new international health regulations. Emerg Infect Dis 2006; 12(7):1058-65.

Balter S, Weiss D, Hanson H, et al. Three years of emergency department gastrointestinal syndromic surveillance in New York City: what have we found? Morb Mort Wkly Rep 2004: 54 Suppl: 175-80.

Berger M, Shiau R, Weintraub JM. Review of syndromic surveillance: implications for waterborne disease detection. J Epidemiol Community Health 2006 Jun; 60(6):543-50.

Bodley-Tickell AT, Mossop H, Rehman Y, Natin D, Blair I. Enhanced surveillance of HIV infection in the West Midlands. Commun Dis Public Health 2004 Dec;7(4):315-8.

Bravata DM, McDonald KM, Smith WM, et al. Systematic review: surveillance systems for early detection of bioterrorism-related diseases. Ann Intern Med 2004 Jun 1;140(11):910-22.

Broome CV. Federal role in early detection preparedness systems. Morb Mort Wkly Rep 2005 Aug 26;54 Suppl:7-9.

Buehler JW, Berkelman RL, Hartley DM, Peters CJ. Syndromic surveillance and bioterrorism-related epidemics. Emerg Infect Dis 2003 Oct;9(10):1197-204.

Buehler, JW, Hopkins RS, Overhage JM, Sosin DM, Tong V. Framework for evaluating public health surveillance systems for early detection outbreaks: recommendations from the CDC Working Group. Morb Mort Wkly Rep 2004; 53 (RR-5):1-11.

Burt J, Hooper R, Jessopp L. The relationship between the use of NHS Direct and deprivation in southeast London: an ecological analysis. J Publ Hlth Med 2003;25: 174-176.

Cartmill M, White BD. Telephone advice for neurosurgical referrals. Who assumes duty of care? Br J Neurosurg 2001; 15(6): 453-455.

Chen JH, Schmit K, Chang H, et al. Use of Medicaid prescription data for syndromic surveillance - New York. Morb Mort Wkly Rep 2005 Aug 26;54 Suppl:31-4.

Chinemana F. Letter to author. July 2004.

Cliff AD, Haggett p, Ord KL, Versey G. Spatial diffusion: An historical Geography of Epidemics in an Island Community. Cambridge: Cambridge University Press 1981. 6-35.

Cooper DL, Arnold E, Smith GE, et al. The effect of deprivation, age and gender on NHS Direct call rates. Br J Gen Pract 2005;5:287-291.

Das D, Metzger K, Heffernan R, et al. Monitoring over-the-counter medication sales for early detection of disease outbreaks - New York City. Morb Mort Wkly Rep 2005; 54 Suppl:41-46.

Department of Health (1997). The New NHS Modern Dependable. The Stationary Office. London.

Department of Health (2002). Getting ahead of the curve: a strategy for combating infectious diseases (including other aspects of health protection). 2002. London: DH.

Department of Health. Hospital Activity Statistics. Department of Health Website. http://www.performance.doh.gov.uk/hospitalactivity/index.htm (accessed 14th April 2008).

Derby MP, McNally J, Ranger-Moore J, et al. Poison Control Center--based syndromic surveillance for foodborne illness. Morb Mort Wkly Rep 2005. 26;54 Suppl:35-40.

Doroshenko A, Cooper D, Smith G, et al. Evaluation of syndromic surveillance based on National Health Service Direct derived data−England and Wales. Morb Mort Wkly Rep 2005; 54 Suppl: 117−22.

Enter-net annual report: 2005 - Surveillance of Enteric Pathogens in Europe and Beyond. http://www.hpa.org.uk/hpa/inter/enter-net/Enter-net%20annual%20report%202005%20final.pdf (accessed 19th March 2008).

Espino JU, Hogan WR, Wagner MM. Telephone triage: a timely data source for surveillance of influenza-like diseases. AMIA Annu Symp Proc 2003:215-219.

European Influenza Surveillance Scheme. http://www.eiss.org/index.cgi (accessed 17th March 2008).

Fleming DM, Smith GE, Charlton JR, et al. Impact of infections on primary care--greater than expected. Commun Dis Public Health. 2002 Mar;5(1):7-12.

Fleming DM. Weekly returns service of the Royal College of General Practitioners. Commun Dis Public Health 1999;2:96-100.

Foster J, Dale J, Jessopp L. A qualitative study of older people's views of out-of-hours services. Br J Gen Pract 2001 Sep;51(470):719-23.

Foster J, Jessopp f, Chakraborti S. Do callers to NHS Direct follow the advice to attend an accident and emergency department? Emerg Med J 2003; 20: 285 - 288.

Gillespie I, Elson R. Successful reduction of human Salmonella Enteritidis infection in England and Wales. Euro Surveill 2005. Nov 17;10(11): E051117.2.

Greatbatch D, Hanlon G, Goode J, et al. Telephone triage, expert systems and clinical expertise. Sociol Health Illn 2005; 27(6):802-30.

Harcourt S, Smith GE, Hippisley-Cox J, et al. Report of the first year of a pilot, national, primary care surveillance project. Proceedings of the Health Protection Agency Annual Conference 2005; p 119.

Hawker J, Begg N, Blair I, et al. Communicable Disease Control Handbook. Oxford: Blackwell Science Limited, 2001; p313.

Health protection Agency (2005). A Surveillance and Response Strategy: Definition, Principles and Overall Strategy. Surveillance Strategy Programme Development Board.

Health Protection Agency (2008). HIV and STIs Surveillance Systems. http://www.hpa.org.uk/webw/HPAweb&Page&HPAwebAutoListName/Page/1202115502759?p=1202115502759 (accessed 22nd April 2008).

Health Protection Agency (2008). Health Protection Report. http://www.hpa.org.uk/webw/HPAweb&HPAwebStandard/HPAweb_C/1203496953186?p=1205914023518 (accessed 14th April 2008).

Heffernan R, Mostashari F, Das D, et al. New York City syndromic surveillance systems. Morb Mort Wkly Rep 2004 Sep 24;53 Suppl:23-7.

Heffernan R, Mostashari F, Das D, et al. Syndromic Surveillance in Public Health Practice, New York City. Emerg Infect Dis 2004; 10(5): 858-864.

Henning KJ. What is syndromic surveillance? Morb Mort Wkly Rep 2004 Sep 24;53 Suppl:5-11.

Henry JV, Magruder S, Snyder M. Comparison of office visit and nurse advice hotline data for syndromic surveillance--Baltimore-Washington, D.C., metropolitan area, 2002. Morb Mort Wkly Rep 2004;53 Suppl:112-116.

Hippisley-Cox J, Smith S, Smith G, et al. QFLU: new influenza monitoring in UK primary care to support pandemic influenza planning. Euro Surveill 2006 Jun 22;11(6):E060622.4.

Hurt−Mullen KJ, Coberly J. Syndromic surveillance on the epidemiologist's desktop: making sense of much data. Morb Mort Wkly Rep 2005; 54 Suppl:141-6.

Hutcherson CM. Legal Considerations for Nurses Practicing in a Telehealth Setting. OJIN 2001: 6;3

Jones NF, Marshall R. Centers for Disease Control and Prevention (CDC). Evaluation of an electronic general-practitioner-based syndromic surveillance system--Auckland, New Zealand, 2000-2001. Morb Mort Wkly Rep 2004 Sep 24;53 Suppl:173-8.

Jones RC, Liberatore M, Fernandez JR, Gerber SI. Use of a prospective space-time scan statistic to prioritize shigellosis case investigations in an urban jurisdiction. Public Health Rep 2006 Mar-Apr;121(2):133-9.

Kaiser R, Coulombier D. Different approaches to gathering epidemic intelligence in Europe. Euro Surveill 2006; 11(4):060427.

Knowles E, Munro J, O'cathain A, Nicholl J. Equity of access to health care. Evidence from NHS Direct in the UK. J Telemed Telecare 2006;12(5):262-5.

Kulldorff M, Heffernan R, Hartman J, Assuncao R, Mostashari F. A space-time permutation scan statistic for disease outbreak detection. PLoS Med 2005 Mar;2(3):e59. Epub 2005 Feb 15.

Lambell P, Cooper A, Hoyles S, Pygall S-A, O'Cathain. An audit of paediatric fever in NHS Direct: consistency of advice by nurses using computerised decision support software systems. Clinical Governance 2003;8:222-226.

Langmuir AD. The surveillance of communicable diseases of national importance. N Engl J Med 1963;268:182–92.

Lattimer V, George S, Thompson F, et al. Safety and effectiveness of nurse telephone consultation in out of hours primary care: randomised controlled trial. BMJ 1998;317:1054–1059.

Lazarus R, Kleinman K, Dashevsky I, DeMaria A Jr, Platt R. Using automated medical records for rapid identification of illness syndromes (syndromic surveillance): the example of lower respiratory infection. BMC Public Health 2001; 1: 9.

Lazarus R, Kleinman K, Dashevsky I, et al. Use of automated ambulatory-care encounter records for detection of acute illness clusters, including potential bioterrorism events. Emerg Infect Dis 2002 Aug;8(8):753-60.

Lombardo JS, Burkom H, Pavlin J. ESSENCE II and the framework for evaluating syndromic surveillance systems. Accession number & update. Morb Mort Wkly Rep 2004; 53 Suppl: 159−65.

Ma H, Rolka H, Mandl K, Buckeridge D, Fleischauer A, Pavlin J. Implementation of laboratory order data in BioSense Early Event Detection and Situation Awareness System. Morb Mort Wkly Rep 2005 Aug 26;54 Suppl:27-30.

Magruder SF, Henry J, Snyde M. Linked analysis for definition of nurse advice line syndrome groups, and comparison to encounters. Morb Mort Wkly Rep 2005;54 Suppl:93-97.

Marquet RL, Bartelds AIM, van Noort SP, et al. BMC Public Health 2006; 6:242.

Marx MA, Rodriguez CV, Greenko J, et al. Diarrheal illness detected through syndromic surveillance after a massive power outage: New York City, August 2003. Am J Public Health 2006;96(3):547-53.

McNulty CA, Smith G. PHLS Primary Care Coordinators. PHLS works towards effective interaction with primary care. Commun Dis Public Health 2000 Dec;3(4):234-9.

Meijer A, Paget WJ, Meerhoff TJ, et al. Epidemiological and virological assessment of influenza activity in Europe, during the 2004-2005 winter. Euro Surv 2006: 11(5).

Micheal D, Pavlin JA, Mansfield LJ et al. Disease outbreak detection system using syndromic data in the greater Washington DC area. Am J Prev Med 2002; 23: 180-186.

Miller B, Kassenborg H, Dunsmuir W, et al. Syndromic surveillance for influenzalike illness in ambulatory care network. Emerg Infect Dis 2004;10(10):1806-11.

Miller E, Farrington CP. The current epidemiology of pertussis in the developed world: UK and West Germany. J Exp Clin Med 1988;13 Suppl:97-101.

Mostashari F, Kulldorff M, Hartman JJ, Miller JR, Kulasekera V. Dead bird clustering: A potential early warning system for West Nile virus activity. Emerg Infect Dis 2001. 9:641-646, 2003.

Munro J, Nicholl J, O'Cathian A, Knowles E. Evaluation of NHS Direct first wave sites: First interim report to the Department of Health. Sheffield: Medical Care Research Unit 1998. University of Sheffield.

Munro J, Nicholl J, O'Cathian, Knowles E, Morgan A. Evaluation of NHS Direct first wave sites: Final report of the phase one research. Sheffield: Medical Care Research Unit 2001. University of Sheffield.

Munro J, Sampson F, Nicholl J. The impact of NHS Direct on the demand for out-of-hours primary and emergency care. Br J Gen Pract 2005; 55(519): 790-2.

Muscatello DJ, Churches T, Kaldor J, et al. An automated, broad-based, near real-time public health surveillance system using presentations to hospital Emergency Departments in New South Wales, Australia. BMC Public Health 2005 Dec 22;5:141.

National Audit Office (2002). Report by the Comptroller and Auditor General: NHS Direct in England. London: The Stationary Office.

Nordin JD, Goodman MJ, Kulldorff M, et al. Simulated anthrax attacks and syndromic surveillance. Emerg Infect Dis 2005 Sep;11(9):1394-8.

O'Cathain A, Nicholl J, Sampson F, et al. Do different types of nurses give different triage decisions in NHS Direct? A mixed methods study. J Health Serv Res Policy 2004 Oct;9(4):226-33.

Osaka K, Takahashi H, Ohyama T. Testing a symptom-based surveillance system at high-profile gatherings as a preparatory measure for bioterrorism. Epidemiol Infect 2002 Dec;129(3):429-34.

Paladini M. Daily Emergency Department Surveillance System, Bergen County, New Jersey. Morb Mort Wkly Rep 2004 Sep 24;53 Suppl:47-9.

Pooley CG, Briggs J, Gatrell A, et al. Contacting your GP when the surgery is closed: issues of location and access. Health and Place 2003; 9(1): 23-32.

Reingold A. If syndromic surveillance is the answer, what is the question? Biosecur Bioterror 2003;1(2):77-81.

Ring F, Jones M. NHS Direct usage in a GP population of children under 5 years: is NHS Direct used by people with the greatest health need? Br J Gen Pract 2004;54:211-213.

Rockx B, van Asten L, van den Wijngaard C, et al. Syndromic surveillance in the Netherlands for the early detection of West Nile virus epidemics. Vector Borne Zoonotic Dis 2006 Summer;6(2):161-9.

Rodman JS, Frost F, Jakubowski W. Using nurse hotlines calls for disease surveillance. Emerg Infect Dis 1998:329-333.

Rolland E, Moore KM, Robinson VA, McGuinness D. Using Ontario's "Telehealth" health telephone helpline as an early-warning system: a study protocol. BMC Health Serv Res 2006;15;6:10.

Rolland E. Comparing Telehealth and NACRS Influenza Activity to Fluwatch Surveillance Using EARS-X. Advances in Disease Surveillane 2007; 4: 189.

Rosen G. A History of Public Health. Baltimore: John Hopkins University Press 1958.

Rowlands S, Moser K. Consultation rates from the general practice research database. Br J Gen Pract 2002; 52: 658-660.

Ruscoe Q, Hill S, Blackmore T, McLean M. An outbreak of Legionella pneumophila suspected to be associated with spa pools on display at a retail store in New Zealand. N Z Med J 2006 Oct 13;119(1243):U2253.

Salisbury C, Trivella M, Bruster S. Demand for and supply of out of hours care from general practitioners in England and Scotland: observational study based on routinely collected data. BMJ 2000 Mar 4;320(7235):618-21.

Salisbury C. The demand for out-of-hours care from GPs: a review. Fam Pract 2000 Aug;17(4):340-7.

Schneider P. Telehealth core to WHO's missions. Healthcare Informatics 1998;July; 59-60.

Shigematsu M, Davison KL, Charlett A, Crowcroft NS. National enhanced surveillance of meningococcal disease in England, Wales and Northern Ireland, January 1999-June 2001. Epidemiol Infect 2002 Dec;129(3):459-70.

Sloane PD, MacFarquhar JK, Sickbert-Bennett E, et al. Syndromic surveillance for emerging infections in office practice using billing data Ann Fam Med 2006 Jul-Aug;4(4):351-8.

Snow J. On the Mode of Communication of Cholera. London: John Churchill 1855.

Sokolow LZ, Grady N, Rolka H, et al. Deciphering data anomalies in BioSense. Morb Mort Wkly Rep 2005 Aug 26;54 Suppl:133-9.

Sosin DM, DeThomasis J. Evaluation challenges for syndromic surveillance--making incremental progress. Morb Mort Wkly Rep 2004 Sep 24;53 Suppl:125-9.

Sosin DM. Syndromic surveillance: the case for skillful investment. Biosecur Bioterror 2003;1(4):247-53.

Steiner-Sichel L, Greenko J, Heffernan R, et al. Field investigations of emergency department syndromic surveillance signals - New York City. Morb Mort Wkly Rep 2004; 53 Suppl:184-9.

Tizzard Z, Edwards JN. Acute respiratory effects following use of waterproofing sprays: Some UK experience. Thorax 2004; 59: 541-a-542-a.

Tran TH, Nguyen TL, Nguyen TD, et al. Avian influenza A (H5N1) in 10 patients in Vietnam. N Engl J Med 2004; 350(12):1171-2.

Turner VF, Bentley PJ, Hodgson SA, et al. Telephone triage in Western Australia. Med J Aust 2002; 176(3):100-3.

Wethington H, Bartlett P.J Usage and data collection patterns for a novel web-based foodborne-disease surveillance system. Environ Health 2006 Mar; 68(7):25-9.

Wheeler JG, Sethi D, Cowden JM, et al. Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. The Infectious Intestinal Disease Study Executive. BMJ 1999 Apr 17; 318(7190):1046-50.

Wilson LE, Smith A, Meyer N, et al. Utilisation of Data Derived from a Nurse-led NHS Access and Information Telephone Helpline (NHS24) in Communicable Disease Management in Scotland. Syndromic Surveillance Conference 2006.

World Health Organisation Executive board (2005). eHealth: proposed tools and services. Report by the Secretariat. December 2005.

Yih K, Teates KS, Abrams A, et al. Nurse Call Data for Detection of Influenza-Like Illness. Advances in Disease Surveillance 2006; 1:76.

Part 1

NHS Direct data and syndromic surveillance

2 NHS Direct derived data: An exciting new opportunity or an epidemiological headache?

Published as: Cooper DL, Chinemana F. NHS Direct derived data: an exciting new opportunity or an epidemiological headache? *J Pub Hlth Med 2004; 26: 158-160.*

Abstract

NHS Direct, a national telephone helpline for health advice, was established in 1998 to provide health information and advice to callers and refer them to an appropriate service. This article briefly describes the nature of the NHS Direct call record and discusses issues relevant to the use of the data for disease surveillance and epidemiological purposes.

Clinical decision support software (NHS CAS) is used by NHS Direct to collect callers' demographic details and direct them to the appropriate level of care. Data relating to NHS Direct calls provide a timely snapshot of symptoms occurring in the community and are summarised in 'off the shelf' NHS CAS reports. Adapting the system to provide customised data extracts requires considerable development work.

When interpreting NHS Direct derived data, particular attention should be given to the age distribution of callers, NHS Direct demand surges, call 'networking' and changes to the NHS CAS clinical algorithms. An increasingly rich source of baseline data, growing body of published work, and a more 'bedded down' NHS Direct service will further our understanding and acceptance of the value of the NHS Direct call record.

Background

A commitment by the Government to provide people with quick and easy access to health care [1] led to the implementation of NHS Direct, a nurse-led national telephone helpline providing confidential healthcare advice [2]. Although few in number, published reviews or applications of NHS Direct data to public health are now raising awareness of this new source of health data [3-5]. The Health Protection Agency Communicable Disease Surveillance Centre has been working in collaboration with NHS Direct since the winter of 1999/2000 [6-9] to explore the use of NHS Direct derived call data for communicable disease surveillance. This article draws on the experience gained during this collaboration and aims to briefly describe how NHS Direct data are collected and stored, discuss issues relevant to the use of NHS Direct data for disease surveillance (and within the wider context of Public Health and epidemiology) and invite opinion from others within the health arena.

The NHS Clinical Assessment System

NHS Direct nurses use clinical decision support software (the NHS Clinical Assessment System (NHS CAS)) to handle calls to NHS Direct. Callers are not diagnosed by NHS Direct but classified ('triaged') on the basis of described symptoms described to determine priority of need and appropriate place of treatment. The NHS CAS is structured around over 200 computerised clinical algorithms. Each algorithm consists of a series of questions relating to the symptoms of the person about whom the call is made. Nurses use clinical judgement to select the most appropriate clinical algorithm and triage the call.

The NHS Direct call record

NHS Direct is organised into 21 sites in England with a single site covering all of Wales. Computerised call records are currently held locally at NHS Direct sites but will eventually be held nationally at one 'data mart'. A call record includes the caller's name, age, sex, address, date of birth, time of call, algorithm (which usually refers to the caller's primary symptom) and disposition (call outcome, e.g. self-care, GP referral, 999 call) as well as a more detailed clinical picture related to the nurses questioning. Ethnicity will be included in the caller's record from summer 2003.

Approximately 70% of all calls answered by NHS Direct are about caller's symptoms and triaged using the NHS CAS. Fifteen to 20% of these calls originate from integrated GP Co-operatives (although this number is as high as 35% at some sites). NHS Direct also handles large numbers of health information queries.

Data availability

The call reporting system linked to the NHS CAS was designed for performance management purposes rather than as an epidemiological tool. However, NHS Direct derived data do provide a valuable snapshot of symptoms in the community and there are a number of 'off the shelf' NHS CAS reports that provide breakdowns of call data (e.g. calls grouped by algorithm or disposition type). Depending on the type of report, call activity can be reported on an hourly, daily, weekly or monthly basis. Adapting the system to provide customised data extracts for surveillance or epidemiological purposes requires further development work.

Although the primary symptom (algorithm) may be the most readily available type of NHS Direct health data, specific symptom based questions, asked by the nurse during the call, could provide the ability to cluster symptoms (e.g. for influenza-like-illness). For example, by following some of the 26 different routes through a 'diarrhoea' algorithm the following groups of symptoms would be identified.

Diarrhoea and vomiting
Diarrhoea and abdominal pain
Diarrhoea and fever
Diarrhoea and recent foreign travel
Diarrhoea and fever and recent foreign travel
Bloody diarrhoea

Unfortunately, the complex structure of decision support algorithms means extracting clusters of symptoms can be technically difficult. The ability to develop syndromic case definitions (clustering groups of symptoms) using NHS Direct call data will be a major factor in determining the extent to which they are used for disease surveillance.

NHS Direct are open to discuss collaborative work with third parties. Negotiation with the NHS Direct National Clinical Team or specific project teams (e.g. the NHS Direct Health Protection Team are currently working with the Health Protection Agency) can provide worthwhile benefits for both organisations. Successful joint working is most likely if the work is first piloted, with ongoing two-way discussion about the validity and appropriateness of call data for the purpose in hand. Regular feedback to NHS Direct sites and their staff ('closing the loop') is essential.

Issues of patient confidentiality must also be carefully considered if details from individual call records are required. A recorded message, heard at the beginning of all calls, implies consent for patient identifiable data (PID) to be used by NHS Direct to improve the service. However, ethics committee approval would be needed for use of PID within the wider NHS or academic and public bodies.

Data quality

Call data should be interpreted with the following caveats.

- At the time of writing NHS Direct call rates are a fraction (approximately one thirtieth) of GP contact rates [8, 10], so may not yet reflect the health of all sections of the population equally.
- Factors such as ethnicity and multiple deprivation are likely to play a large part in determining local awareness and uptake of the service [11, 12].
- NHS Direct attained national coverage in four waves between 1998 and 2000. Since 1999 a number of GP co-operatives have integrated with local NHS Direct sites. Some of the geographical variation in NHS Direct call rates may be explained by these two factors.
- During times of peak demand or staff training, NHS Direct re-routes (networks) local calls to other call centres. Calls originating from a discrete geographical area could potentially be handled by two or more call centres. Therefore, call data from a single NHS Direct site may not truly reflect local demand for the service. This issue is particularly relevant when mapping local variations in call data. Analyses of NHS Direct call data should ideally be done by postcode of residence rather than NHS Direct site.

- Commuting in and out of urban areas may modify NHS Direct catchment populations on an hourly, daily and weekly basis. Adults phoning NHS Direct from work rather than home may cause an increase in urban call rates which will not reflect the true demand on local health services.
- There are surges in demand for NHS Direct during weekends, bank holidays (particularly over Christmas), advertising campaigns, particular health incidents (e.g. over 5000 calls were taken in January 2001 regarding the Alder Hey organs incident) and during NHS Direct integration with local GP co-operatives (usually during non-surgery hours).
- The age distribution of NHS calls is heavily skewed towards children, similar to the age distribution of GP consultations [10]. A third of all NHS Direct symptomatic calls are about children under 15 and a quarter about those aged under 4 years [8]. The elderly are less likely than young adults to phone NHS Direct.
- Regular upgrades to the NHS CAS software mean algorithms may be removed, merged or added to the system. This affects baseline data.
- Relatively short historical baseline data are available from which to interpret call trends.

Future considerations

NHS Direct data can be acquired in a timely manner, in a consistent format across England and Wales and cover a wide variety of symptoms. However, many factors cause surges in demand for NHS Direct. When making temporal or geographical comparisons of symptom based NHS Direct data, careful consideration must be made in choosing an appropriate indicator to explain call patterns (e.g. call numbers, call rates or proportions of calls).

The public and medical communities are not as familiar with NHS Direct data as with more established sources of health data (e.g. Hospital Episode Statistics). Therefore, when presenting NHS Direct data a description of the data (‘metadata’) and any necessary caveats should be included.

Despite limitations, the value of NHS Direct derived data for public health purposes is likely to increase due to:

- More robust call baselines.
- An increasing body of published work.
- The NHS Direct technical infrastructure and flow of information becoming more ‘bedded down’.
- An increasingly experienced NHS Direct staff and analytical capacity.
- Closer integration with main stream public health networks (e.g within PCTs [13]) which may enhance opportunities for joint working.
- The expansion of NHS Direct call taking capacity over the next 3-4 years linked to further integration with out-of-hours primary care providers [13], which will increase NHS Direct call volumes and may improve the representativeness of call data.

Of future significance is the face to face software for the NHS CAS. This software is already used in NHS Walk-in-Centres and some A&E departments and may provide scope to address more definitive diagnoses. This will also provide opportunities to link patient data from multiple health care settings and may provide further insight into the effect NHS Direct has had on other health services [14] and surveillance systems[15].

The answer to the question posed by the title of this article is undoubtedly ‘both’, and can be viewed in a positive light. Careful use of NHS Direct derived data can only further our understanding and acceptance of its value.

References

1. Department of Health. The New NHS Modern and Dependable. Department of Health. 1997
2. NHS Direct Online. About NHS Direct. http://www.nhsdirect.nhs.uk/about/
3. 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.
4. 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, July 2001.
5. Payne F, Jessopp L. NHS Direct: review of activity data for the first year of operation at one site. *J Pub He Med* 2001; 23(2): 155-158.
6. Harcourt S, Smith GE, Hollyoak V et al. Can calls to NHS Direct be used for Syndromic Surveillance? A pilot at three sites using influenza as an example. *Commun Dis Public Health*. 2001; 4(3): 178-182
7. Cooper DL, Smith GE, Hollyoak VA, Joseph CA, Johnson L, Chaloner R. Use of NHS Direct calls for the surveillance of influenza – a second year's experience. *Commun Dis Public Health 2002;* 5(2*):* 132-136*.*
8. Cooper DL, Smith GE, O'Brien SJ, Hollyoak V, Long S. What can Analysis of Calls to NHS Direct Tell us about the Epidemiology of Gastrointestinal Infections in the Community? J of Infection 2003; 46(2): 101-105.
9. Asgari-Jirhandeh N, Lightfoot N, Kramer J, Cooper DL, Gregor S, Collins E, Nicoll A. On behalf of PHLS Deliberate Release Response Teams. PHLS response to the threat of bioterrorism in England & Wales. *International Conference on Emerging Infectious Diseases*. Conference Proceedings 2002. Atlanta. P78.
10. Rowlands S, Moser K. Consultation rates from the general practice research database *Br J Gen Pract*. 2002 Aug; 52(481): 658-660.
11. The Comptroller and Auditor General. *NHS Direct in England.* London: Stationery Office, 2002.
12. Burt J, Hooper R, Jessopp L. The relationship between the use of NHS Direct and deprivation in southeast London: an ecological analysis. *J Pub He Med* 2003; 25(2): 174-176.
13. Department of Health. Developing NHS Direct: A strategy document for the next 3 years. 2003.
14. Munro J, Nicholl JP, O'Cathain A, Knowles E. Impact of NHS Direct on demand for immediate care: observational study. *BMJ* 2000; 321:150-153.
15. 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: 1397-1398.

3 The effect of deprivation, age and gender on NHS Direct call rates

Published as: Cooper DL, Arnold E, Smith GE, Hollyoak VA, Chinemana F, Baker M and O'Brien SJ. The effect of deprivation, age and gender on NHS Direct call rates. *Br J Gen Pract 2005; 5: 287-291*.

Abstract

Background: The use of Primary Care services in the UK is traditionally high in deprived areas. There has been little research into the effect of deprivation on uptake of NHS Direct, a national nurse led health helpline.

Aim: To explore the impact of deprivation, age and sex on call rates to two NHS Direct sites.

Design of study: Ecological study.

Setting: West Yorkshire and West Midlands NHS Direct sites.

Methods: Details of NHS Direct calls between July 2001 and January 2002 were linked to electoral wards and the 'indices of multiple deprivation 2000'. Age standardised call rates were calculated for five deprivation levels. Using a negative binomial regression model, West Yorkshire call rates were analysed by age group, sex, deprivation level and geographical location. Rates were mapped by ward for West Yorkshire NHS Direct.

Results: Six monthly call rates were highest for children under 5 years (130 per 1,000 population). The rate ratio of female to male calls (all ages) was 1.30 (95%CI: 1.27-1.33), this difference being most pronounced in 15-44 year olds (p<0.001). For both West Yorkshire and West Midlands NHS Direct call rates (all ages combined) were highest in areas within the middle of the range of deprivation. West Yorkshire call rates about those under 5 years were lower in the most deprived areas than in the least deprived areas (<1 year: p=0.06, 1-4 years: p=0.03). For adults of 15-64 years, call rates were significantly higher in the most deprived areas (p<0.001).

Conclusion: This work supports previous research in that overall demand for NHS Direct is highest in areas where deprivation is at or just above the national average. This study additionally suggests that the effect of extreme deprivation appears to raise adult call rates but reduce rates about children.

HOW THIS FITS IN

What do we know?

The use of Primary Care services in the UK is traditionally high in deprived areas.
At one NHS Direct site, call rates were shown to rise with increasing deprivation before declining at extreme levels of deprivation. There has been no published work additionally exploring the impact of age and sex, as well as deprivation, on the use of NHS Direct.

What does this paper add?

This work supports previous work in that call rates (all ages combined) are highest in areas with deprivation levels at, or just above, the national average. It additionally suggests that the effect of extreme deprivation appears to raise adult call rates but reduce rates about children.

Introduction

NHS Direct is a national 24-hour health initiative that began operation in 1998 [1]. Since 1999, NHS Direct derived call data have been used for the surveillance of symptoms which may be indicative of infection [2, 3]. These call data indicate that, although there is national coverage by NHS Direct, call rates are still a fraction (approximately one thirtieth) of GP consultation rates [4-6]. NHS Direct calls may therefore not be representative of the population of England and Wales.

Questionnaire studies have indicated low awareness [7] and low usage [8] of NHS Direct amongst more deprived social groups. However, low awareness does not necessarily indicate low usage, which for other primary care services tends to be high in deprived areas [4, 9-11]. A single published study, using a large sample of NHS Direct's own call data, has examined the relationship between deprivation and the use of NHS Direct [12], and challenged previous assumptions that NHS Direct is under utilised in deprived areas [13]. Work at other NHS Direct sites has found that the distribution of NHS Direct calls by age and sex is similar to that for GP services [14, 15]. We aimed to build on previous work by conducting an ecological analysis of call rates at two NHS Direct sites, to assess the effect of deprivation on call rates and to additionally explore the impact of age and sex.

Methods

The age, sex and full postcode of residence of all NHS Direct 'triage' calls (calls relating to symptoms) were collected from West Yorkshire NHS Direct and West Midlands NHS Direct for a 6 month period (July 2001 to January 2002). Records were excluded if the age or postcode were incomplete or missing. Permission to collect these data was obtained from the West Midlands Multi-centre Research Ethics Committee.

To help manage demand NHS Direct sites often 'network' calls to NHS Direct sites elsewhere in the country. Calls may therefore not originate from the catchment area (based on area dialling codes) of the site that handled the call. For the purposes of our study, proxy NHS Direct catchment areas were aggregated from 1998 Health Authority boundaries. Calls were categorised as networked calls, and excluded from the study, if the postcode was located outside the site catchment area.

The ward level 'Index of Multiple Deprivation 2000' score (IMD) was used as a measure of deprivation [16]. All wards in England were ranked according to their IMD score and divided into five deprivation quintiles, each quintile comprising 20% of the population of England. A higher IMD score or deprivation quintile indicates increasing deprivation. A Geographical Information System (ArcView GIS) was used to link each NHS Direct call record to a ward and hence a deprivation quintile. The 1998 mid-term population estimates for wards in England were obtained [17]. For both NHS Direct sites, six monthly age standardised call rates per 1000 population (all ages) were calculated for each deprivation quintile. We also calculated the female to male call rate ratio (all ages (age standardised)) for West Yorkshire NHS Direct.

To explore the impact of age, sex and deprivation on West Yorkshire call rates we used a negative binomial regression model (which can account for over dispersion of call rates). The numbers of calls for each ward were calculated for the following age groups; <1, 1-4, 5-14, 15-44, 45-64 and ≥65 years (males and females separately). The ward population of each age-sex category was calculated using 1991 Census population data, adjusted using 1998 mid-term population estimates for wards. Due to the imprecise nature of the West Yorkshire NHS Direct boundary (constructed from Health Authorities boundaries), we felt call rates around the periphery of our catchment area could be low due to cross boundary flows of calls. To account for this the GIS was used to categorise each West Yorkshire ward into either an 'edge ward' (adjacent to the boundary: n=24) or an 'inner ward' (enclosed within the edge wards: n=102). The outcome variable for our model was call rate with independent variables being sex, age-group, deprivation quintile and edge ward. Interactions between variables were tested for significance. The negative binomial regression model was not used for West Midlands call data as the sex of the caller was missing from a substantial number of records.

Six monthly call rates per 1,000 population were mapped by Ward for West Yorkshire NHS Direct.

Results

There were 57,662 calls to West Yorkshire NHS Direct and 52,446 to West Midlands NHS Direct of which 40,345 (70%) for West Yorkshire and 45,156 (86%) for West Midlands were included in the study. The majority of excluded calls were networked calls. After exclusions full postcode was available for more than 98% of call records, at both sites. The six monthly call rates to West Yorkshire and West Midlands NHS Direct, after exclusions, were 19.1 and 19.8 calls per 1000 population respectively.

Deprivation

Neither the West Yorkshire (population 2.2 million) nor West Midlands NHS Direct (2.4 million) catchment areas conform to the national deprivation profile. These two sites have 35% and 45% of their catchment population within the most deprived 20% of the population of England (Table 1).

Table 1. Ranges of index of multiple deprivation scores; percentage of the population in each deprivation quintile; and call rate ratio by deprivation quintile (all ages (age standardised)) for West Yorkshire NHS Direct and West Midlands NHS Direct. The baseline quintile is quintile 1 (least deprived).

West Yorkshire NHS Direct					West Midlands NHS Direct			
Range of scores	Deprivation quintile	Percentage of the population	Rate ratio	Confidence interval (95%)	Deprivation quintile	Percentage of the population	Rate ratio	Confidence interval (95%)
1.16-10.16	1	6%	1.00	-	1	9%	1.00	-
10.16-16.82	2	21%	1.27	(1.20-1.33)	2	10%	1.31	(1.25-1.37)
16.82-26.06	3	20%	1.45	(1.38-1.53)	3	13%	1.41	(1.36-1.47)
26.06-39.92	4	18%	1.50	(1.43-1.58)	4	23%	1.40	(1.35-1.45)
39.97-83.77	5	35%	1.31	(1.25-1.38)	5	45%	1.22	(1.18-1.27)

For both West Yorkshire and West Midlands NHS Direct call rates (all ages combined) rose with increasing deprivation before falling in the most deprived areas (Table 1). The highest call rates were in quintiles 3 and 4 for both sites (those areas with deprivation equal or just above the national average). Call rates were significantly lower in the least deprived wards (quintile 1) than any other areas (quintiles 2 to 5).

Regression model

There were significant interactions within the model between age-group and deprivation (p<0.001) and between age-group and sex (p<0.001) (see following sections). Edge ward was not associated with these interactions and does not bias the results of this study.

Age and sex

The highest 6-monthly call rates to West Yorkshire NHS Direct were for children aged under one year (130 calls per 1,000) and between one and four years (58 per

1000) (Table 2). The rate ratio of female to male calls (all ages) was 1.30 (95%CI: 1.27-1.33). Call rates were higher about boys than girls although this difference was not statistically significant (Table 2). Rates were higher about women than men, this difference being most pronounced for 15-44 year olds ($p<0.001$).

Table 2. Proportions of calls; six monthly call rates per 1000 population; and female to male call rate ratios by age group for West Yorkshire NHS Direct.

Age band	Proportion of calls (%)	6 monthly call rate per 1000 population	Female: Male rate ratio (95%CI)	*p*-value
<1 year	6.6	130.3	0.93 (0.85-1.03)	0.16
1-4 years	17.3	58.4	0.93 (0.86-1.02)	0.10
5-14 years	11.7	17.3	0.98 (0.89-1.07)	0.60
15-44 years	42.6	18.7	1.99 (1.84-2.15)	<0.001
45-64 years	11.8	10.7	1.57 (1.43-1.72)	<0.001
≥65 years	6.9	8.5	1.15 (1.03-1.28)	0.01

Deprivation, age and sex

For West Yorkshire NHS Direct, call rates for the <1 year and 1-4 year age groups were lowest in the most deprived wards (<1 year: p=0.06, 1-4 years: p=0.03), significantly so for the 1-4 year age group (Table 3). For adults of working age (15-64 years), call rates rose with increasing deprivation and were significantly higher in the most deprived areas (quintiles 3-5) than in the least deprived. There was no significant variation in call rates for older people (≥65 years). The relationship between call rates and deprivation was the same for males and females as it was for both sexes combined (i.e this relationship did not differ significantly between men and women (data not shown)).

Location

Call rates for edge wards (the periphery of the West Yorkshire catchment area) were 5.1% lower than for inner wards but this difference was not significant (rate ratio of edge to inner wards = 0.95; 95%CI: 0.90-1.00). Six monthly call rates for West Yorkshire wards varied between 11.3 and 42.4 per 1000 population (median = 24) and were highest in and around the Halifax and Huddersfield urban areas and South of Leeds (Figure 1).

Discussion

Summary of main findings

The impact of deprivation on NHS Direct call rates in West Yorkshire differs between age groups. The effect of extreme deprivation appears to raise adult call rates but reduce rates about children. When considering total demand for the service (all ages), call rates to West Midlands and West Yorkshire NHS Direct were highest in areas with deprivation levels at or just above the national average. The relationship between call rates and deprivation was the same for males, females and both sexes combined.

Strengths and limitations of the study

This study has used adjusted population data to calculate call rates. Population data and deprivation indices generated from the 2001 census will improve the accuracy of this type of study in the future. Some bias may have occurred as a result of not including calls made within our two NHS Direct site catchment areas but networked elsewhere in England. This bias is considered minimal as calls are not networked on the basis of age or relative location within a site boundary. The use of a national call database, from which calls can be extracted by postcode, could resolve this problem. The non-inclusion of networked calls also means our analysis will underestimate the true call rates. This study was able to eliminate any bias caused by the imprecise nature of our NHS Direct site boundary ('edge effects') by distinguishing between inner and edge wards. Finally, the local distribution of primary care services, other than NHS Direct, may explain some of the four-fold variation in call rates between wards in West Yorkshire. Although methodologically difficult, we recommend that proximity to other health services, as well as the ethnic composition of the catchment areas, should be taken into account when analysing call rates for further studies. This could help clarify how existing health services, as well as social factors, have influenced the relative uptake of NHS Direct.

Comparison with existing literature

These call data support previous work [15, 18] in that the distribution of calls by age and sex is comparable to that for GP services [4-5, 9-10], apart from the low call rate

Table 3. Call rate ratios (RR) for West Yorkshire NHS Direct by age-group and deprivation quintile. The baseline quintile is quintile 1 (least deprived).

Age group	<1 year		1-4 years		5-14 years		15-44 years		45-64 years		≥65 years	
Quintile	RR	(95% CI); *p*-value	RR	(95% CI); *p*-value	RR	(95% CI); *p*-value	RR	(95% CI); *p*-value	RR	(95% CI); *p*-value	RR	(95% CI); *p*-value
1	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-
2	1.02	(0.8-1.3); 0.88	1.06	(0.85-1.32); 0.61	1.24	(0.98-1.59); 0.7	1.26	(1.03-1.54); 0.02	1.28	(1.02-1.62); 0.03	1.07	(0.83-1.37); 0.60
3	1.16	(0.91-1.48); 0.22	1.28	(1.03-1.59); 0.03	1.63	(1.29-2.08); <0.0001	1.40	(1.15-1.71); <0.001	1.49	(1.19-1.89); <0.001	1.12	(0.87-1.43); 0.38
4	1.09	(0.85-1.39); 0.48	1.10	(0.87-1.37); 0.38	1.61	(1.27-2.06); <0.0001	1.61	(1.31-1.97); <0.0001	1.43	(1.13-1.81); 0.003	1.16	(0.91-1.5); 0.23
5	0.80	(0.63-1.01); 0.06	0.80	(0.65-0.99); 0.03	1.32	(1.05-1.67); 0.02	1.56	(1.29-1.89); <0.0001	1.48	(1.18-1.85); <0.001	0.97	(0.76-1.23); 0.78

from those over 65 years. What is more interesting is how the effect of deprivation on NHS Direct call rates is different from its effect on the use of other primary care services. Previous work has demonstrated high GP consultation rates from low social classes for adults [4], and children [9], and high demand for out of hours doctor contact in deprived areas [10-11]. Conversely, the pattern of NHS Direct call rates at three NHS Direct sites (West Midlands, West Yorkshire and previous work in SE London [12]) indicates a drop in call rates between average and severely deprived areas. The analysis of West Yorkshire call data adds that this drop does not occur for all age groups. Call rates about young children were lowest in deprived areas where as call rates about adults were generally highest in deprived areas. The distribution of local health services, behavioural factors, and some bias due to the ecological fallacy [19] could explain this.

Implications for further research

The IMD score is a recently produced index and has been used as a proxy for deprivation on an individual level, as opposed to the more commonly used Jarman index [20] and Townsend score [21], which are based on 1991 Census data. The IMD score is timely as it comprises 1991 Census data as well as government socio-economic data from the late 1990s. The use of national IMD quintiles in this study, rather than site specific quintiles, has helped to establish some baseline data for comparison with similar work at other NHS Direct sites.

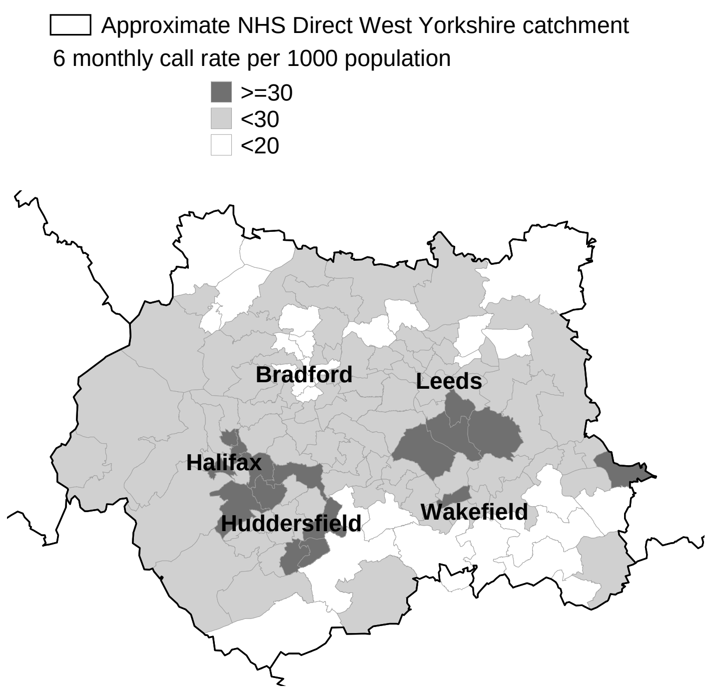
In this study we analysed calls from people with symptoms where as a fifth of NHS Direct calls are from people not reporting symptoms but seeking health information. We are unsure how call rates for health information vary by age, sex and local deprivation. Two sites with predominantly urban populations were studied here and further research is required to examine whether NHS Direct has been successful in reaching rural communities. Provisional analysis of call data from one site (Midlands Shires NHS Direct) suggests call rates are significantly lower in rural areas than in urban areas [22].

NHS Direct call profiles, presented here, will be most useful if regular updates are produced. They may then be of use to NHS Direct to assess demand for the service and evaluate the success of promotional activities. Analysis of national call data should now be undertaken to determine whether call rates about children in deprived areas are also low in other parts of the country.

Conclusion

This work supports previous research in that NHS Direct call rates are highest in areas where deprivation levels are at, or just above, the national average [12]. This study additionally suggests that the effect of extreme deprivation appears to raise adult call rates but reduce rates about children. Work is needed to examine whether this relationship exists across the country and on an individual level. The significant local variations in call rates demonstrated here (by age, sex and deprivation) should be considered by NHS Direct when developing the service, and by public health staff when using call data for health protection and surveillance purposes [3].

Figure 1. Map of 6 monthly call rates per 1000 population by ward for West Yorkshire NHS Direct.



References

1. NHS Direct online. http://www.nhsdirect.nhs.uk/
2. Cooper DL, Smith GE, Hollyoak VA, Joseph CA, Johnson L, Chaloner R. Use of NHS Direct calls for the surveillance of influenza – a second year's experience. *Commun Dis Public Health* 2002; **5(2)**: 132-136.
3. Baker M, Smith GE, Cooper DL, Verlander NQ, Chinemana F, Cotterill S, Hollyoak V, Griffiths R. Early Warning and NHS Direct: a role in community surveillance? J Publ Hlth Med 2004; **25**:362-368.
4. Royal college of General Practitioners, Office of Population, Censuses and Surveys, Department of Health. Morbidity Statistics from General Practice: Fourth National Study, 1991-1992. London. HMDSO, 1995.
5. Rowlands S, Moser K. Consultation rates from the general practice research database. *Br J Gen Pract* 2002*;* **52**: 658-660.
6. Directorate of Access and Choice. Developing NHS Direct: A strategy document for the next 3 years. London, Department of Health, 2003.
7. The Comptroller and Auditor General. NHS Direct in England. London: Stationery Office, 2002.
8. Ring F, Jones M. NHS Direct usage in a GP population of children under 5 years: is NHS Direct used by people with the greatest health need? *Br J Gen Pract* 2004; **54**: 211-213.
9. Saxena S, Majeed A, Jones M. Socioeconomic differences in childhood consultation rates in general practice in England and Wales: prospective cohort study. *BMJ* 1999; **318**: 642-6.
10. Salisbury C, Trivella M, Bruster S. Demand for and supply of out of hours care from general practitioners in England and Scotland: observational study based on routinely collected data. *BMJ* 2000; **320**: 618 - 621.
11. Carlisle R**,** Groom LM, Avery AJ, Boot D, Earwicker S. Relation of out of hours activity by general practice and accident and emergency services with deprivation in Nottingham: longitudinal survey. *BMJ* 1998; **316:** 520 - 523.
12. Burt J, Hooper R, Jessopp L. The relationship between the use of NHS Direct and deprivation in southeast London: an ecological analysis. *J Publ Hlth Med* 2003; **25**: 174-176.
13. George S. NHS Direct audited. *BMJ* 2002; **324**: 558-559.
14. 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, 1999.
15. 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.
16. Department of Environment, Transport and the Regions (2000). 'Indices of Deprivation.' Regeneration Research Summary No. 37, DETR, London.
17. Penhale B, Noble M, Smith G, Wright G. Final consultation on ward level population estimates. Social Disadvantage Research Group 1999. http://www.odpm.gov.uk/stellent/groups/odpm_urbanpolicy/documents/page/odpm_urbpol_608095.pdf.
18. Payne F, Jessopp L. NHS Direct: review of activity data for the first year of operation at one site. *J Publ Hlth Med* 2001; **23**: 155-158.
19. Robinson WR. Ecological correlations and the behaviour of individuals. American Sociological Reviews 1950; **15**: 351-557.
20. Jarman B. Identification of Under-Privileged Areas. *BMJ* 1983; **286**: 1705-1709.
21. Townsend P. (1979).'Poverty in the UK.' Penguin, London.
22. Cooper D. The socio-demographic profile of NHS Direct callers – A study at 3 NHS Direct sites. Provisional results – September 2002.

4 National Symptom Surveillance Using Calls to a Telephone Health Advice Service — United Kingdom, December 2001–February 2003

Published as: Cooper DL, Smith GE, Baker M, Chinemana F, Verlander N, Gerard E, Hollyoak V, Griffiths R. National Symptom Surveillance Using Calls to a Telephone Health Advice Service — United Kingdom, December 2001–February 2003. *MMWR 2004; 53 (Supp): 179-183.*

Abstract

Introduction: Recent terrorist activity has highlighted the need to improve surveillance systems for the early detection of chemical or biological attacks. A new national surveillance system currently examines symptoms reported to a telephone health advice service in the UK.

Objectives: The aim is to identify an increase in symptoms indicative of the early stages of illness caused by the deliberate release of a biological or chemical agent or more common infections.

Methods: Data relating to 10 key symptoms/syndromes (mainly respiratory and gastrointestinal) are received electronically from 23 call centres covering England and Wales. Data are analysed on a daily basis and statistically significant excesses ('exceedances') in calls are automatically highlighted and assessed by a multi-disciplinary team.

Results: Between December 2001 and February 2003 there were 1811 exceedances of which 126 required further investigation and 16 resulted in 'alerts' to local or national health protection teams. Examples of these investigations are described.

Conclusion: Surveillance of NHS Direct calls has detected high levels of activity in specific symptoms at both national and regional level. Although no deliberate release has so far been detected, this surveillance system continues to be refined.

Introduction

Recent terrorist activity has highlighted the need to improve surveillance systems for the early detection of chemical or biological attacks. A new UK surveillance system currently examines symptoms reported to a national telephone health advice service (NHS Direct) [1]. NHS Direct is a nurse-led helpline which aims to provide the public with rapid access to professional health advice and information about health, illness and the National Health Service (NHS) [2]. The service is open 365 days a year and covers the entire population of England and Wales. Calls are priced at local call rates. NHS Direct nurses use clinical decision support software, the NHS Clinical Assessment System (NHS CAS), to respond to calls. The NHS CAS contains over 200 clinical algorithms which form tree like structures of questions relating to the symptoms of the person about whom the call is made. Most calls result in a call outcome; either advice for self-care; a routine doctor referral; an urgent doctor referral; an Accident and Emergency Department referral; or a paramedic callout. It has been suggested that NHS Direct derived data could be of value in the surveillance of disease.

Where a deliberate release of a harmful agent causes illnesses with an extended mild prodromal phase, a proportion of people are likely to contact NHS Direct in the first instance. Consequently there may be an opportunity to identify an increase in illness before it is identified by other primary care or secondary care services. The aim of the surveillance described here is to identify an increase in symptoms indicative of the early stages of illness caused by the deliberate release of a biological or chemical agent, or illness caused by common infections. This builds on existing surveillance of influenza-like-illness and gastrointestinal symptoms using NHS Direct calls [3-5].

Method

Daily call data relating to the following 10 algorithms (symptoms/syndromes) are received electronically by the Health Protection Agency (HPA) from all 23 NHS Direct sites in England and Wales: cold/flu, cough, diarrhoea, difficulty breathing, double vision, eye problems, lumps, fever, rash, vomiting. The surveillance of eye problems calls replaced ‘food poisoning’ calls in April 2003. These data are analysed on a daily basis by a surveillance team, established in November 2001 and consisting of HPA and NHS Direct staff. The 10 symptoms/syndromes are intended to be indicative of the early stages of a range of illness caused by biological or chemical weapons. Data are broken down by NHS Direct site, symptom, age-group and call outcome. NHS Direct nurses triage rather than diagnose illness in callers.

Upper confidence limits (99.5% level) of calls for each symptom, as a proportion of daily total calls, are constructed for each NHS Direct site. These confidence limits are derived from standard formula for proportions [6] with the baseline numbers of total calls and symptom calls adjusted for seasonal effects (winter: December to February, spring: March to May, summer: June to August and autumn: September to October). When the daily proportion of calls is above the 99.5% upper confidence limit this is termed an 'exceedance'.

In addition to the confidence interval analyses, control charts are constructed for five of the ten symptoms/syndromes (cold/flu, cough, fever, diarrhoea, vomiting) at the 10 NHS Direct sites covering major urban centres (London, Manchester, Leeds, Birmingham and Newcastle). Baselines for the control charts are calculated by assuming the number of syndromic calls follow a Poisson distribution with the total number of calls as an offset. A model is fitted to each site and symptom separately using data from December 2001. The model always contains a public holiday and seasonal term. If shown to be necessary, a day of the week (weekday, Saturday or Sunday) and a linear long-term trend factor are also fitted. Scaling is performed to account for over-dispersion when this is present.

A normal approximation is not used to calculate the 99.5% upper control chart limit of calls for each symptom as it yields more exceedances than would be expected (i.e. >0.5%). Alternatively, a transformation to approximate normality with zero mean is performed and back-transformed to the original scale. For control charts, the following formula for the 99.5% upper limit of syndromic calls is used:

$$\left(\sinh\left(\frac{z_{\alpha}/2 + \sqrt{N-0.5}\,\sinh^{-1}\sqrt{p}}{\sqrt{N-0.5}}\right) \right)^2 (N-0.75) - 3/8$$

where N is given by the expected value divided by one less than the scale parameter, p is equal to the scale parameter minus one and z_α is the 100*(1-α)th centile of the Normal distribution. Ad-hoc choices of z are made to achieve the desired number of purely random exceedances (0.5%). The upper 99.5% control chart limit of calls for each symptom, as a proportion of total calls, is calculated on a daily basis.

Statistically significant excesses (termed 'exceedances') in calls for any of the 10 symptoms/syndromes are automatically highlighted (for the confidence interval and control chart method) and assessed by the surveillance team (stage 1). If no reasonable explanation for the exceedance can be found additional line listings of call details (including the call ID and caller's home postcode (similar to US zipcodes)) are requested for the day of the exceedance and the current date (stage 2). The call ID, which should be a unique number, is used to identify duplicate call records. Requesting calls for the current date (which will be complete up to the hour the request is made) is considered important to monitor what may be an evolving situation. If current call data indicate persistently high calls for a particular algorithm a geographical information system (GIS) can be used to map call data, although this is not a routine procedure for all exceedances.

NHS Direct is a 'virtual organisation' in the sense that individual NHS Direct sites may 'network' (export) calls to other sites during periods of peak demand. A proportion of calls handled by NHS Direct sites (usually less than 10%) may therefore originate from outside their catchment areas. Catchment areas are based on local area telephone dialling codes.

If the surveillance team believe the information provided by the line listings (stage 2) necessitates further investigation they then pass call information to the relevant local or national public health teams for follow-up by existing mechanisms (termed 'alerts': stage 3). Also, if the exceedance is thought to represent a serious public health threat, the NHS Direct Medical Adviser can contact callers to obtain further clinical information. Weekly bulletins, summarising NHS Direct call activity, are disseminated to relevant local and national health protection colleagues.

Table 1. Numbers of exceedances (Stage 1), exceedances investigated (Stage 2) and 'Alerts' (Stage 3) by algorithm (December 2001 – February 2003).

Syndrome	*Stage 1* Number of exceedances *	*Stage 2* Exceedances investigated	*Stage 3* Alerts†
Fever	328	23	4
Cough	279	4	0
Cold/influenza	185	5	0
Vomiting	182	28	4
Double vision	180	2	0
Food poisoning	180	0	0
Lumps	142	14	1
Diarrhoea	137	14	4
Difficulty breathing	123	22	2
Rash	75	13	1
TOTAL	1,811	126	16

* An exceedance is a statistically significant excess of calls beyond the 99.5% upper-confidence limit.

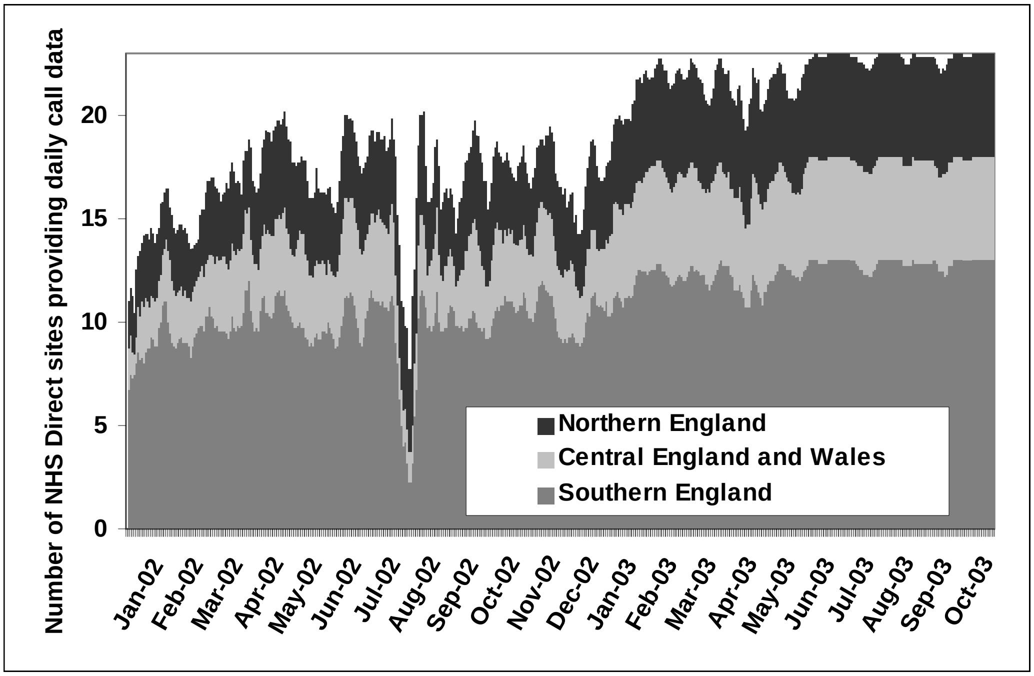
† Refer to earlier published work for a description of Stage 3 exceedances (*1*).

Results

When the surveillance of 10 symptoms/syndromes began in December 2001 call data were collected from approximately half of the total 23 NHS Direct sites. Subsequently, the mean number of NHS Direct sites providing daily call data rose from 12 in December 2001 to all 23 during October 2003 (Figure 1). A sudden drop in the number of sites providing call data occurred in July 2002 due to surveillance staff absences. There are no consistent differences between regions in the proportion of sites providing data.

Between December 2001 and February 2003 there were 1811 confidence limit exceedances (stage 1) (Table 1), of which 126 (7%) required further investigation (stage 2), and 16 (1%) resulted in alerts (stage 3). The main reasons for not progressing from an exceedance investigation (stage 2) to an alert (stage 3) were because the observed rise in calls was a single day exceedance only (46% of stage 2 investigations), duplicate call records caused the exceedance (20%) and the lack of geographical clustering of the call data (15%).

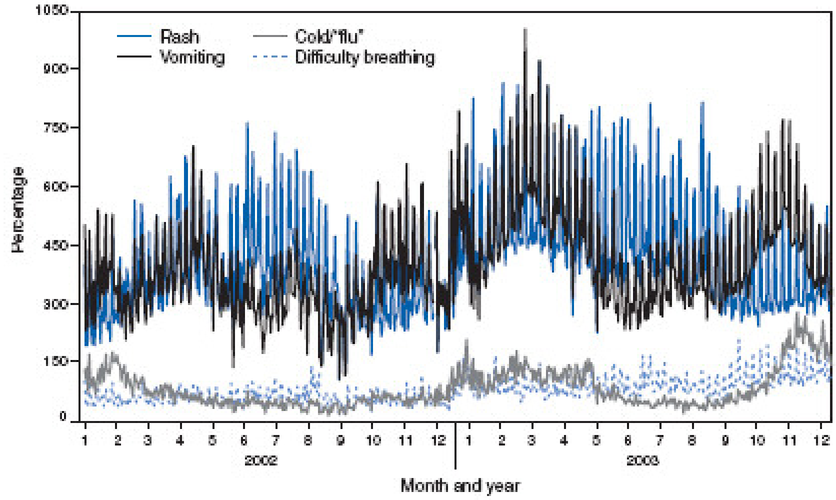
Figure 1. Number of NHS Direct sites providing daily call data in Southern England, Central England and Wales, and Northern England (December 2001 – October 2003), expressed as a 7 day rolling average.



Figures 2 and 3 give an overview of the national daily number of calls and daily proportions of calls for four algorithms. As expected, a seasonal pattern of higher activity over the winter is emerging in some of the algorithms (e.g. 'cold/flu' and vomiting), both in the numbers and proportions of calls. The numbers of calls for all 10 algorithms rises at weekends and during public holidays, when many routine primary care services are closed. The proportion of calls in some algorithms also rises at weekends (e.g. rash) and during public holidays (e.g. cough, vomiting).

During early August 2002 there were a number of daily exceedances in difficulty breathing calls within the Thames basin and East Anglia. These exceedances accompanied a general rise in difficulty breathing calls in eastern parts of Central and Southern England (Figure 4). Prior to the rise there were high ozone levels and thunderstorms in this part of England. We are currently analysing the timing and impact of these climatic and environmental conditions on call data. Also, as a consequence of detecting this sudden rise in calls, which may be linked to environmental conditions, we are forging operational links with experts in environmental health.

Figure 2. National daily numbers of NHS Direct calls for ‘cold/flu’, difficulty breathing, vomiting and rash (This graph includes back data collected retrospectively for July 2002).



* Includes back data collected retrospectively for July 2002.

Figure 3. National daily proportions of NHS Direct calls for ‘cold/flu’, difficulty breathing, vomiting and rash.

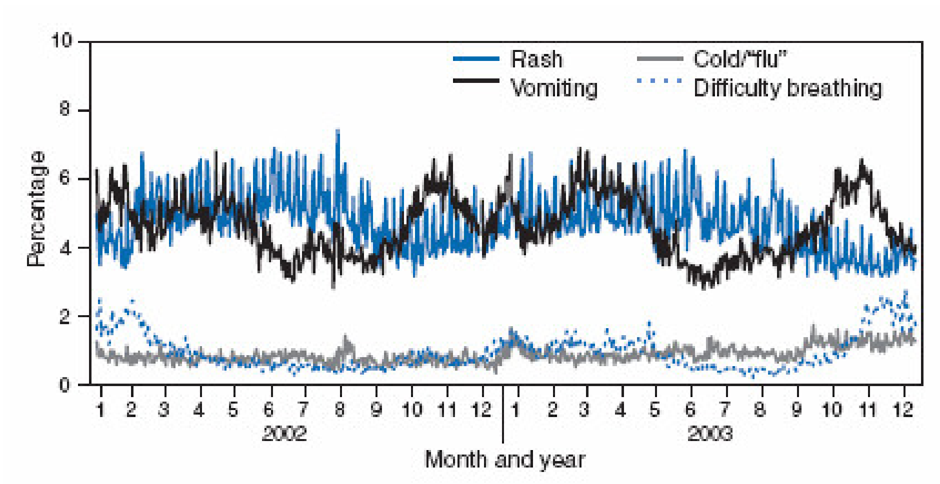
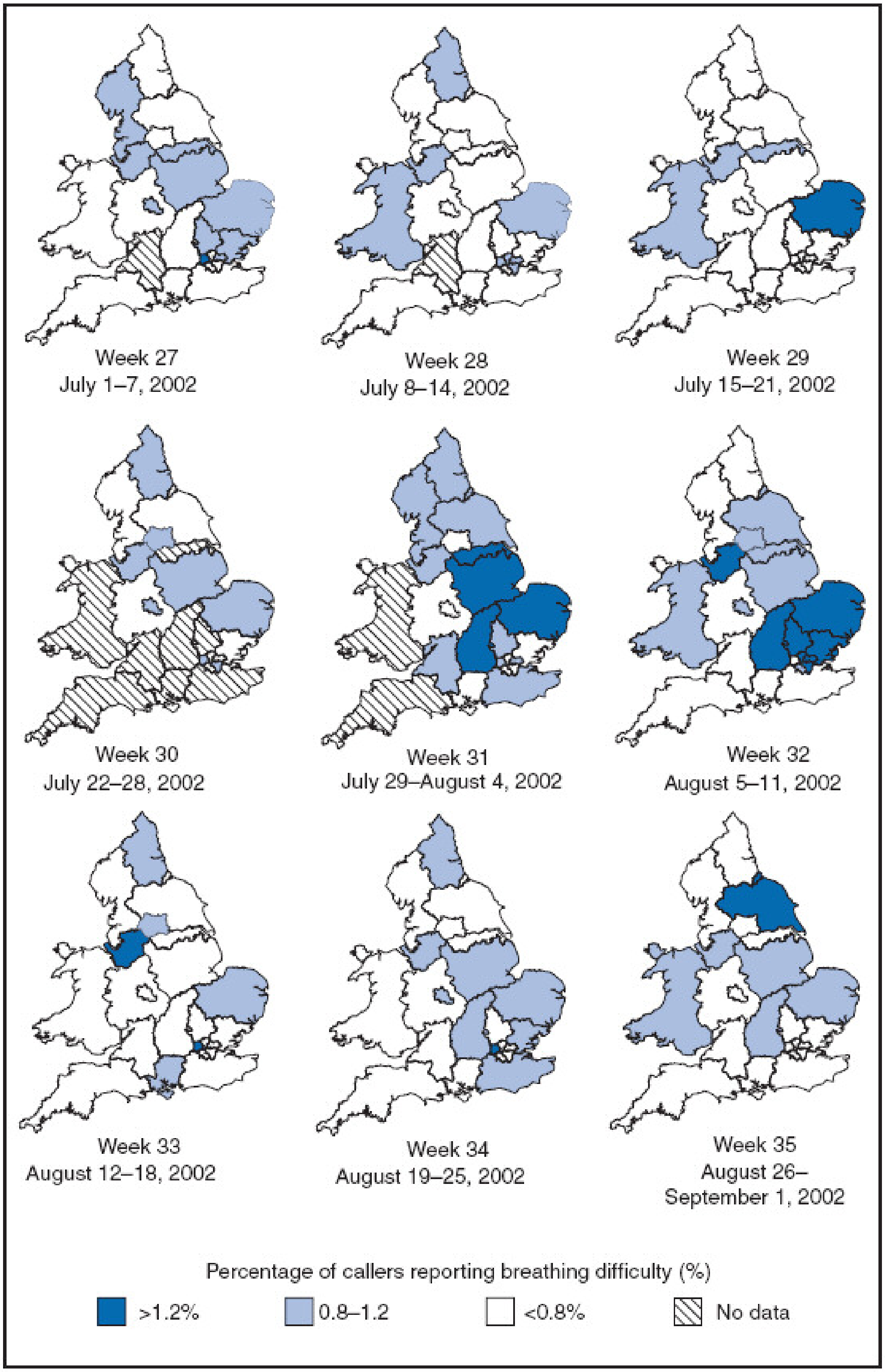


Figure 4. Increase in the proportion of difficulty breathing calls in NHS Direct sites in Eastern England (July/August 2002).



In January 2003, traces of the chemical poison ricin were found in a North London apartment. As a response to this incident the surveillance team was requested to enhance symptom surveillance of call data collected from the 5 NHS Direct sites in London. Data were collected on a number of symptoms/syndromes (Figure 5) and updated every 2 hours. Call data were also mapped, by place of residence rather than by workplace, as this may have provided the first clues that a deliberate release could have occurred at a particular point source. Experience so far from NHS Direct data and other data sources have demonstrated no evidence of any deliberate release of biological or chemical agents within the UK.

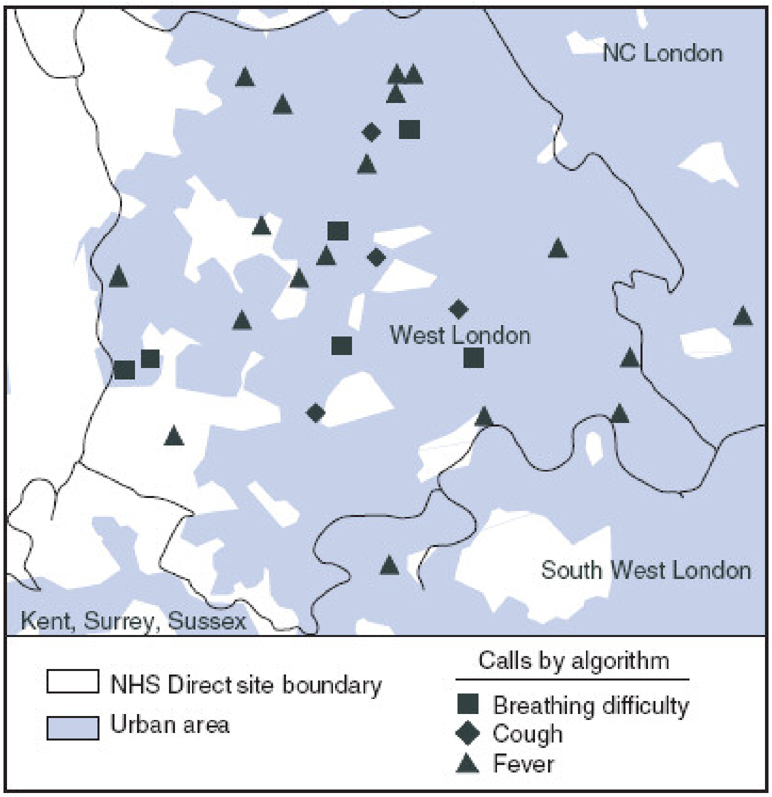
Conclusions

This is the only syndromic surveillance system covering all the population of England and Wales. Although the majority of exceedances do not result in any further action being taken, when action is taken, health protection teams are normally informed within 24-48 hours of the calls to NHS Direct being made. At the time of writing, only two years data have been collected and the establishment of baselines and refinement of the statistical methodology is ongoing. Although no deliberate release of chemical or biological agents has been detected, this surveillance system has detected high levels of activity in specific symptoms at both national and regional levels.

After an initial period when duplicate call records led to the investigation of exceedances that proved to be spurious, improvements in data quality have been made. The surveillance now covers the entire population of England and Wales and is conducted on a daily basis. Although the geographical locations of calls are available on request, the geographical resolution of the initial daily analysis (to identify exceedances) is at a site level. This means localised sub-site level outbreaks could be missed. The surveillance team is currently investigating ways to collect and analyse call data by smaller geographical units.

Consistent and timely data returns have been achieved through concentrating on collecting routine NHS Direct data, attempting to cause minimal disruption to the work patterns of the data providers (NHS Direct sites) and ensuring continual feedback to all staff within the surveillance network. The annual running cost of the current surveillance system is under $200,000. Importantly, the provision of these surveillance data is now an integral part of NHS Direct's objectives and maintains a high priority within the service. This was essential when, in January 2003, real time surveillance was needed to address a perceived increased threat of a deliberate release. In this instance analysis if the surveillance data provided some reassurance that a deliberate release had not occurred.

Figure 5. Calls to West London NHS Direct for Difficulty Breathing, Cold/flu, Cough and Fever (7th of January 2003), mapped by home postcode.



A recent Government strategy document has announced a three fold expansion of NHS Direct call handling capacity over the next 3-4 years, with an equivalent increase in call volumes also expected (from 6 to 16 million calls per year in England) [7]. This compares with approximately 14 million visits a year to Accident and Emergency Departments in England [8] and 190 million consultations with primary care doctors [9]. The increase in NHS Direct call volumes should improve the representativeness of the call data and potential for early identification of disease outbreaks.

The value of NHS Direct derived surveillance data, in complementing existing surveillance for common infections (e.g. influenza) is becoming more established [1,3,4]. It is too early to say that the NHS Direct symptom surveillance system will ultimately provide an early warning of a chemical or bio-terrorist attack. Only future events will demonstrate whether this is indeed the case.

References

1. Baker M, Smith GE, Cooper D, Verlander NQ, Chinamena Frances, Cotterill S, Hollyoak V, Griffiths R. Early Warning and NHS Direct: a role in community surveillance? *J Pub H Med* 2003: 24(4); 362-368.
2. NHS Direct online. http://www.nhsdirect.nhs.uk/
3. Harcourt S, Smith GE, Hollyoak V, Joseph CA, Chaloner R, Rehman Y, Warburton F, Ejidokun OO, Watson JM, Griffiths RK. Can calls to NHS Direct be used for Syndromic Surveillance? A pilot at three sites using influenza as an example. *Commun Dis Public Health*. 2001: 4(3); 178-82.
4. Cooper DL, Smith GE, Hollyoak VA, Joseph CA, Johnson L, Chaloner R. Use of NHS Direct calls for the surveillance of influenza – a second year's experience. C*ommun Dis Public Health 2002; 5(2):* 132-136.
5. Cooper DL, Smith GE, O'Brien SJ, Hollyoak V, Long S. What can analysis of calls to NHS Direct Tell us about the Epidemiology of Gastrointestinal Infections in the community? *J of Infection* 2003; 46(2): 101-105.
6. Armitage P, Berry G. *Statistical method in medical research (2nd Edition)*. Oxford, Blackwell Scientific Productions, 1987.
7. Directorate of Access and Choice. Developing NHS Direct: A strategy document for the next three years. Dept of Health, London, April 2003.
8. Hospital Activity Statistics. Department of Health Website. http://www.performance.doh.gov.uk/hospitalactivity/index.htm.
9. Rowlands S, Moser K. Consultation rates from the general practice research database. *Br J Gen Pract* 2002*;* 52: 658-660.

Part 2

Evaluation studies - external validity and comparison against other data sources

5 Can syndromic surveillance data detect local outbreaks of communicable disease? A model using a historical cryptosporidiosis outbreak

Published as: Cooper DL, Verlander NQ, Smith GE, Charlett A, Gerard E, Willocks L, O'Brien S. Can syndromic surveillance data detect local outbreaks of communicable disease? A model using a historical cryptosporidiosis outbreak. *Epidemiol Infect 2006; 134 (1): 13-20.*

Abstract

A national UK surveillance system currently uses data from a health helpline (NHS Direct) in an attempt to provide early warning of a bio-terrorist attack, or an outbreak caused by a more common infection. To test this syndromic surveillance system we superimposed data from a historical outbreak of cryptosporidiosis onto a statistical model of NHS Direct call data. We modelled whether calls about diarrhoea (a proxy for cryptosporidiosis) exceeded a statistical threshold, thus alerting the surveillance team to the outbreak. On the date that the public health team were first notified of the outbreak our model predicted a 4% chance of detection when we assumed one-twentieth of cryptosporidiosis cases telephoned the helpline. This rose to a 72% chance when we assumed nine-tenths of cases telephoned. The NHS Direct surveillance system is currently unlikely to detect an event similar to the cryptosporidiosis outbreak used here and may be most suited to detecting more widespread rises in syndromes in the community, as previously demonstrated. However, the expected rise in NHS Direct call rates, should improve early warning of outbreaks using call data.

Introduction

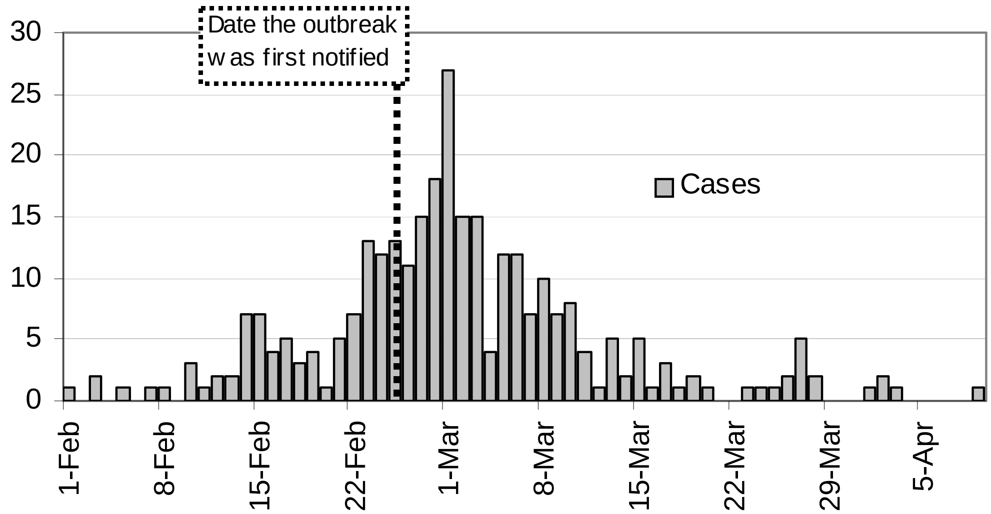
In the UK, laboratory and clinical data form the basis of community surveillance of infectious diseases [1-2]. Since the winter of 1999/2000 NHS Direct derived call data have been used to monitor community levels of syndromes which may be related to infection [3-4]. NHS Direct is a 24 hour nurse led helpline covering the whole of England and Wales [5]. Nurses use the NHS Clinical Assessment System (NHS CAS) which is based around 200 clinical algorithms. Initially work focussed on the use of NHS Direct data for the surveillance of influenza-like-illness, but was expanded in November 2001 as a response to the terrorist and anthrax attacks in the US. National call data relating to 10 key symptoms/syndromes (algorithms) are collected from all 23 NHS Direct sites and analysed on a daily basis by a surveillance team. The aim is to identify increases in syndromes which may represent the prodromal stages of disease caused by a bio-terrorist attack, or more likely a rise in common infections.

To date no bio-terrorist attack has been detected in the UK, or elsewhere in the world [6-9], using syndromic surveillance systems. Our surveillance system has therefore never been subjected to this challenge. Although it has been possible to identify rises in calls for certain syndromes, we need to gain an understanding of what disease outbreaks we are likely to detect. Therefore we have superimposed data from a real outbreak of cryptosporidiosis onto the current NHS Direct surveillance system. *Cryptosporidium* is considered a good example to use due to the relatively long incubation period, which may lead to an insidious rise in cases in the community [1]. We aimed to determine whether a historical cryptosporidiosis outbreak would have been detected using the current NHS Direct surveillance system and, if so, at what point during the outbreak it would have been detected.

Methods

Case data were obtained from a cryptosporidiosis outbreak that occurred during the spring of 1997 in North London and Hertfordshire [10]. The dates of onset of the 345 laboratory confirmed cases varied between February 1 and April 9 and peaked on March 1 (Figure 1). The outbreak was first notified on the 25 of February. Drinking unboiled tap water from a treatment works was the cause of the outbreak.

Figure 1. Epidemic curve of cryptosporidiosis cases from the London outbreak (February – April 1997).

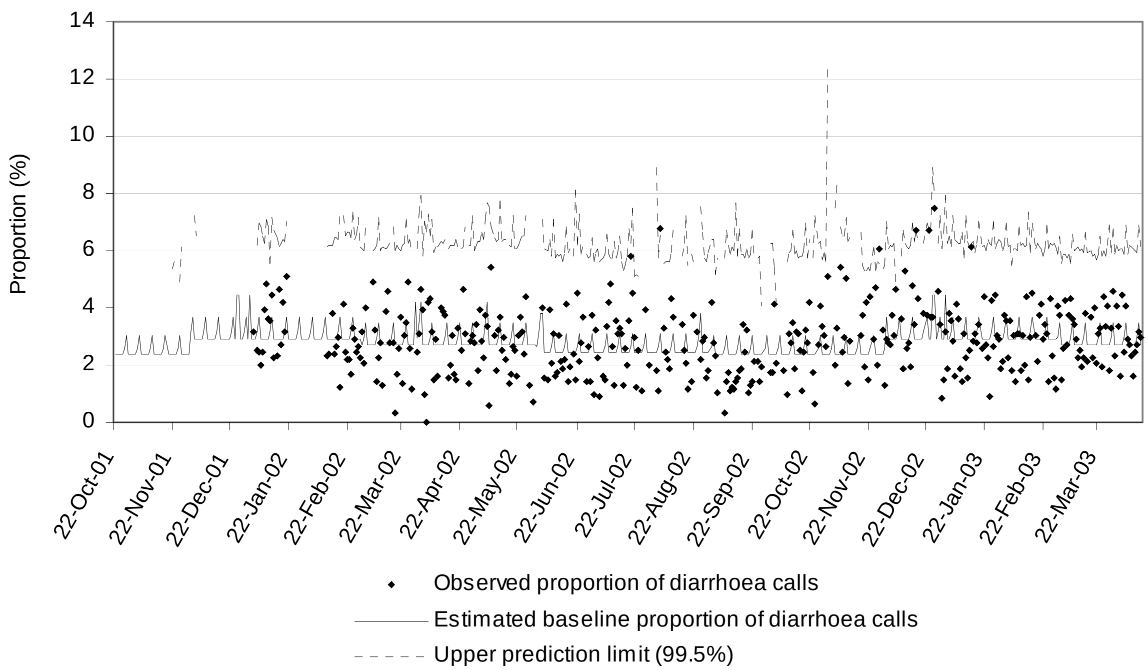


We assumed that cases of *cryptosporidium* would report diarrhoea. Data on total daily calls and calls where the diarrhoea algorithm was used were obtained from NHS Direct North Central London (approximate population: 1.6 million) for the 22nd of October 2001 to the 14 of April 2003. Observed diarrhoea calls, as a proportion of total calls, are displayed in Figure 2. A baseline number of daily NHS Direct diarrhoea calls was estimated, assuming a Poisson distribution, incorporating a bank holiday factor, seasonal factor and day factor (weekday, Saturday or Sunday) (Figure 2). Within our surveillance system we use 2 methods to calculate upper prediction limits for the proportion of diarrhoea calls. The first method, termed the 'control chart method', uses the estimated baseline number of diarrhoea calls to calculate a daily 99.5% upper prediction limit for the proportion of diarrhoea calls (Figure 2). The second method, termed the 'confidence interval method', calculates 99.5% upper confidence intervals derived from standard formula for proportions [11]. Both these methods are described in the Annexe to this report.

The fraction of cryptosporidiosis cases who would have telephoned NHS Direct is unknown so was estimated by first taking the known number of cryptosporidiosis cases which were microbiologically confirmed (n). We know from a national infectious intestinal disease study [12] that approximately half of the cases of cryptosporidiosis seen by the GP will be reported to a laboratory, and therefore estimate that 2n cases will have been seen by their GP. The total call rate to NHS

Direct is approximately one fortieth of the consultation rate of GPs, this proportion being derived from comparing the NHS Direct total call rate to the consultation rate for all diseases and conditions reported to GPs [13-14]. We have therefore assumed that the fraction of community cases of cryptosporidiosis that telephone NHS Direct is about one fortieth of the GP cases. For this outbreak this would be 2n/40 = n/20. Thus taking the laboratory confirmed cryptosporidiosis cases in this outbreak (n), we estimate that 1/20 (one twentieth) of these cases would have telephoned NHS Direct. (Due to the uncertainty of this fraction we also modelled one-hundredth, one fiftieth, one-tenth, three-tenths, six-tenths and nine-tenths of cases telephoning NHS Direct). We assumed that cases would telephone NHS Direct on the first day of their symptoms.

Figure 2. Control chart showing (in the absence of the added outbreak data) the observed proportion of diarrhoea calls; the simulated baseline number of diarrhoea calls (expressed as a proportion of total calls); and the simulated 99. 5% upper prediction limit.



We simulated a daily number of diarrhoea calls from our estimated baseline (Annexe). An outbreak observation of diarrhoea calls (‘cryptosporidiosis’) was then generated from a Poisson distribution with mean given by the fraction of cryptosporidiosis cases telephoning NHS Direct. The outbreak number of diarrhoea calls was added to the simulated daily number of diarrhoea calls. The total number of calls on each day was also increased by these additional calls. When the proportion of diarrhoea calls (simulated plus outbreak calls) exceeded the upper 99.5% control chart limit or upper 99.5% confidence interval this was termed an exceedance. Ten thousand such simulations were conducted. Days with no outbreak cases or where NHS Direct data were unavailable were omitted from the analysis. For the simulations, we assumed that the cryptosporidiosis outbreak occurred in the same period in 2003 (February to April).

The median and maximum proportion of exceedances in the absence of the cryptosporidiosis outbreak were calculated for both methods.

For both methods (control chart method and confidence interval method) we calculated the maximum and median number of single day exceedances (out of 10,000) for each day, for the range of fractions modelled. The maximum percentage excess above the upper prediction limits on each day and the number of occasions where an exceedance one day was followed by an exceedance the next day (a pair of exceedances) were calculated. Time series graphs were plotted showing the likelihood of detecting a single day exceedance on any given day, for the range of fractions of cases ringing NHS Direct. The results for the one-hundredth, one fiftieth and one-tenth fractions are not presented in this report due to their similarity to the results for the one-twentieth fraction.

Results

Between 22 of October 2001 to the 14 of April 2003 North Central London NHS Direct handled 24,516 calls of which 738 were classified as diarrhoea calls (3% of the total). The range of daily total calls and daily diarrhoea calls was 206 to 508 and 3 to 22 respectively. The range of daily diarrhoea calls as a proportion of total calls was 1.1% to 4.6% with a median of 3%.

Model accuracy

Using the 99.5% control chart limit the median proportion of exceedances, in the absence of outbreak data, was 0.48% and the maximum 0.74%. This is close to the expected 0.5%. For the 99.5% confidence interval method (second method used) the median error rate was 2.1% and the maximum 5.7%, both much higher than the expected 0.5% (due to ignoring factors such as bank holidays and over-dispersion of the call data).

Control chart method

Number and size of exceedances: When the outbreak data were added to the model, for the control chart method the maximum number of exceedances out of 10,000 simulations ranged from 102 (1%) when one-twentieth of cryptosporidiosis cases telephoned NHS Direct to 9,651 (97%) when nine-tenths of cases telephoned (Table 1). The maximum number of exceedances occurred on the 1 March for all fractions of cases telephoning NHS Direct. The median percentage excess from 10,000 simulations ranged from 32% when one-twentieth of cases telephoned to 51% when nine-tenths telephoned (Table 1).

Likelihood of observing an exceedance: The likelihood of observing a single day exceedance remained below 20% (for all fractions of cases phoning NHS Direct) until the 23rd of February. The likelihood then began to rise, peaking on the 1st of March (Figure 3). A smaller early peak in the likelihood of observing an exceedance occurred on the 14th of February, caused by a small early peak in cryptosporidiosis cases on the same date. On the date the outbreak was first notified (25 February), our model predicted a 1% (if one-twentieth of cases telephoned NHS Direct) to 47% chance (if nine-tenths of cases telephoned NHS Direct) of detecting the outbreak.

Successive (two day) exceedances: The first successive exceedances, for the range of cases phoning NHS Direct, occurred between 8 and 15 February, prior to the date the outbreak was first notified (25 February).

Confidence interval method

Number and size of exceedances: For the upper confidence interval method, the maximum number of exceedances out of 10,000 ranged from 729 (7%) when one-twentieth of cryptosporidiosis cases telephoned NHS Direct to 9,908 (99%) when nine-tenths of cases phoned (Table 1). The maximum number of exceedances occurred on the 14 of February for the one-twentieth fraction of cases, and on the 1 March for all other fractions**.** The median percentage excess from 10,000 simulations ranged from 54% when one-twentieth of cases telephoned to 72% when nine-tenths phoned.

Table 1. Median, maximum, and date of maximum numbers of single day exceedances out of 10,000 simulations; median, maximum, and date of maximum percentage excess above the upper prediction limit from 10,000 simulations; date of first successive (two day) exceedance, maximum successive exceedances, and date of maximum successive exceedances out of 10,000 simulations; for all fractions of *cryptosporidium* cases telephoning NHS Direct (one-twentieth, three-tenths, six-tenths, nine-tenths).

		Fraction of *cryptosporidium* cases telephoning NHS Direct			
		one-twentieth	**three-tenths**	**six-tenths**	**nine-tenths**
Single of day exceedances out of 10,000 simulations					
Control chart method	Median number of single day exceedances	57.5	100	180	289
	Maximum number of single day exceedances	102	1839	6923	9651
	Date of maximum number of single day exceedances	01/03/2003	01/03/2003	01/03/2003	01/03/2003
Confidence interval method	Median number of single day exceedances	247.5	386	644	1073.5
	Maximum number of single day exceedances	729	3149	8469	9908
	Date of maximum number of single day exceedances	14/02/2003	01/03/2003	01/03/2003	01/03/2003
Percentage excess above the upper prediction limit from 10,000 simulations					
Control chart method	Median percentage excess above the upper prediction limit	32%	38%	49%	51%
	Maximum percentage excess above the upper prediction limit	57%	78%	102%	134%
	Date of maximum percentage excess above the upper prediction limit	14/02/2003	27/02/2003	27/02/2003	01/03/2003
Confidence interval method	Median percentage excess above the upper prediction limit	54%	59%	67%	72%
	Maximum percentage excess above the upper prediction limit	94%	120%	152%	161%
	Date of maximum percentage excess above the upper prediction limit	13/03/2003	14/02/2003	14/02/2003	01/03/2003
Successive (two day) exceedances out of 10,000 simulations					
Control chart method	Date of 1st successive exceedances	15/02/2003	3/02/2003	8/02/2003	1/02/2003
	Maximum number of successive exceedances	2	60	620	363
	Date of maximum number of successive exceedances	04/03/2003 06/03/2003	1/03/2003	1/03/2003	1/03/2003
Confidence interval method	Date of 1st successive exceedances	08/02/2003	8/02/2003	8/02/2003	8/02/2003
	Maximum number of successive exceedances	21	01	022	841
	Date of maximum number of successive exceedances	28/03/2003	1/03/2003	1/03/2003	1/03/2003

Likelihood of observing an exceedance: A similar pattern to the control chart method was observed (Figure 4), although the likelihood of observing exceedances was higher throughout the study period. When the outbreak was first notified (25 February), our model predicted a 4% (if one-twentieth of cases called) to 72% chance (if nine-tenths of cases called) of detecting the outbreak.

Successive (two day) exceedances: The first successive exceedances, for the range of cases telephoning NHS Direct, occurred on the 8 February, again prior to the date the outbreak was first notified.

Figure 3. Likelihood of detecting a single day exceedance on a given day during the outbreak (control chart method). Each curve represents an estimate of the fraction of cryptosporidiosis cases who telephoned NHS Direct.

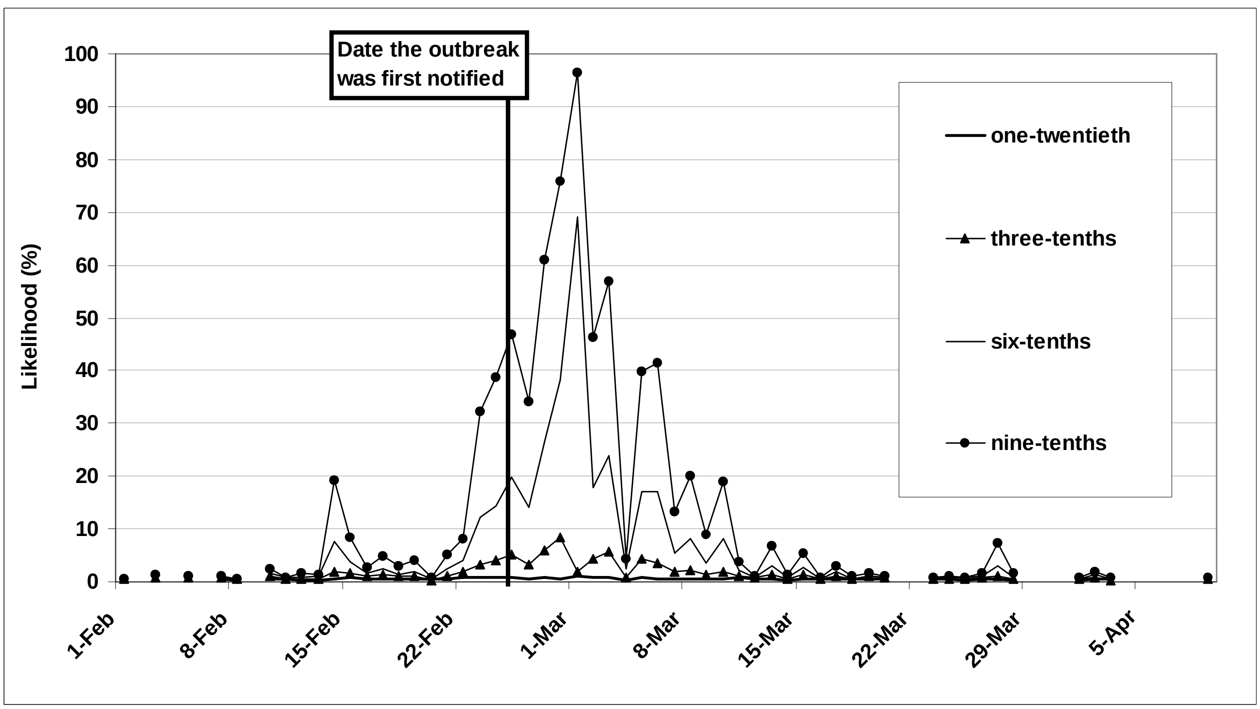
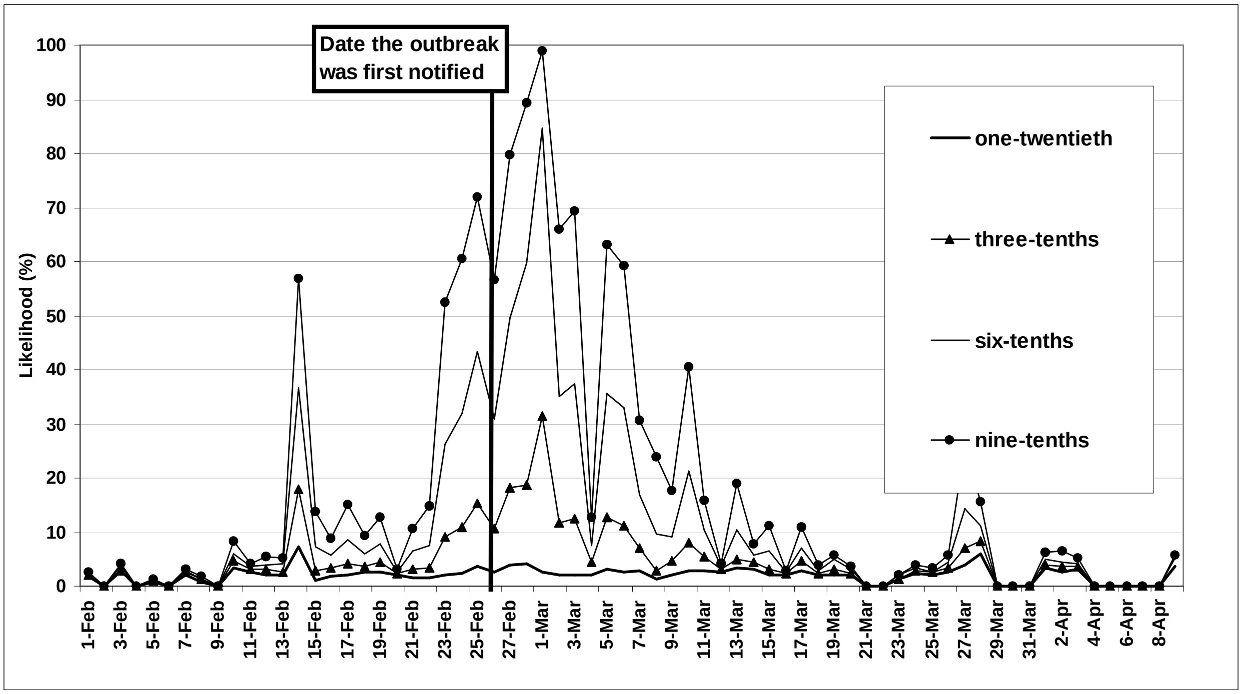


Figure 4. Likelihood of detecting a single day exceedance on a given day (confidence interval method). Each curve represents an estimate of the fraction of cryptosporidiosis cases who telephoned NHS Direct.



Discussion

The NHS Direct syndromic surveillance system is currently unlikely to provide early detection of an event similar to the cryptosporidiosis outbreak modelled here. The system may be most suited to detecting widespread generalised rises in syndromes in the community. This has been previously demonstrated [3-4]. However, our model has shown that as the fraction of cryptosporidiosis cases telephoning NHS Direct increases, the chance of detecting the outbreak increases. When nine-tenths of cases telephoned NHS Direct (using the confidence interval method) there was a one in two chance of detection prior to the date officials were notified of this outbreak. Successive day exceedances may currently be of little practical use as they occur infrequently, unless we assume a very high fraction of cases telephone NHS Direct (nine-tenths). In practice the surveillance team also uses other factors in deciding whether to fully investigate an exceedance, such as seasonality, other surveillance data, call activity at neighbouring NHS Direct sites and the severity of call outcomes.

The median daily excess above the prediction limits when one-twentieth of cases telephone NHS Direct (our current estimate), was 32% for the control chart method and 54% for the confidence interval method. This is a positive result as both these levels of excess are likely to trigger further investigation of the exceedance by the surveillance team. More detailed call information would be obtained (e.g. geographical location of calls), possibly leading to a timely intervention by local health protection teams. Of the two methods of detecting exceedances described here (control chart and confidence interval methods), the confidence interval method produces more exceedances throughout the outbreak period but also more false positives. This is mainly due to over-dispersion of the data, which is accounted for within the more rigorous control chart method.

Various assumptions have been made for this work. To test our surveillance system we have used a single outbreak of cryptosporidiosis, modelled using a single proxy syndrome (diarrhoea), at a single NHS Direct site. We do not know what fraction of *cryptosporidium* cases would have telephoned NHS Direct. Our conservative estimate of outbreak related diarrhoea calls (17 for the one-twentieth estimate), has been superimposed on a relatively high total of diarrhoea calls (738 during the outbreak period). The most commonly reported symptoms for this outbreak were diarrhoea, abdominal pain, fever and vomiting [10]. We assumed that all cryptosporidiosis cases would have been dealt with using the NHS Direct diarrhoea algorithm whereas some may have been dealt with using other algorithms. This would have diluted a rise in calls between two or more algorithms, possibly resulting in no exceedances for any single algorithm. NHS Direct uses call networking (transfer of calls between sites) to manage national demand for the service. We also assumed that no call networking, either in or out of North Central London NHS Direct, occurred during our study period.

The prediction limits used here may be too high for the current system to detect outbreaks of this kind. Strict criteria, 99.5% limits, have been used as opposed to the more commonly used 95% limits. When using the control chart method, exceedances early in the outbreak were rare. There is therefore an argument for lowering the control chart prediction limits (e.g. to 95% or 98%), in an attempt to increase the likelihood of early detection. This would increase the number of exceedances whilst maintaining the scientific validity of using control charts (i.e. incorporating important factors in the model). Lowering exceedance limits will need to be balanced against an acceptable increase in false positives (‘naturally occurring exceedances’). City wide syndromic surveillance systems used in the US [8] and Japan [9] have reported using 95% significance limits. These studies did not report an excess amount of false positives although it is not clear from the reports whether these data also exhibited over-dispersion. 'Multiple models' have been suggested as a means of improving the sensitivity of syndromic surveillance data for outbreak detection [15]. These models have the advantage of adjusting for naturally occurring outbreaks (e.g. influenza) in an attempt to detect simultaneously occurring unnatural outbreaks (e.g. resulting from a bio-terrorist attack). 'Multiday temporal filters' [16], as opposed to the one and two day exceedances modelled here, have also improved detection sensitivity under simulated outbreak conditions.

We do not yet know if the full potential for the system to act as an early warning system has been reached. Assuming a high proportion of cryptosporidiosis cases will telephone NHS Direct may be unrealistic given current NHS Direct call rates. However, a recent government paper has outlined plans for expanding NHS Direct, with a threefold increase in calls predicted by 2007 [17]. A corresponding increase in the number of outbreak cases telephoning NHS Direct should improve our surveillance tool for outbreak detection. It is worth noting that syndromic surveillance is a relatively new field. There is of yet no consensus of opinion, originating from the different systems around the world, as to the best outbreak detection methods and what type of outbreaks syndromic surveillance will or will not detect. Individual systems are likely to become more targeted at specific types of events [18] (either naturally occurring or due to a deliberate release) as the body of evaluation work increases. Due to its national population coverage, and an emphasis on prodromal illness (health calls), further evaluation of the NHS Direct surveillance system should enhance this body of work.

In conclusion, given our present assumptions, the NHS Direct surveillance system is currently unlikely to detect an event similar to this cryptosporidiosis outbreak. Significant daily rises in NHS Direct diarrhoea calls did occur early in the outbreak but were low in number. We know that the system detects national rises in syndromes which coincide with rises detected by other surveillance systems (e.g. influenza). The full potential of the surveillance system for detecting local outbreaks is likely to be reached only after a substantial rise in call rates. This work has provided useful information about the nature and volume of syndromic data required to trigger ‘exceedances’ (signals). Work now needs to focus on defining a statistical threshold that will maximise the detection of true rises in syndromes, whilst avoiding naturally occurring fluctuations. Further modelling work is also required to determine the size and nature of outbreaks that this surveillance system is most likely to detect.

References

1. Noah N, O’Mahony, M. Communicable Disease Epidemiology and Control. Chichester: John Wiley and Sons Ltd, 1998: Chap 10.
2. Hawker J, Begg N, Blair I, Reintjes R, Weinberg J. Communicable Disease Control Handbook. Oxford: Blackwell Science Limited, 2001: Section 4.2.
3. Baker M, Smith GE, Cooper D, et al. Early Warning and NHS Direct: a role in community surveillance? J Pub H Med 2003; 25: 362-368.
4. Cooper DL, Smith GE, Hollyoak VA, et al. Use of NHS Direct calls for the surveillance of influenza – a second year’s experience. Commun Dis Public Health 2002; 5: 132-136.
5. NHS Direct Online. http://www.nhsdirect.nhs.uk/
6. Ackelsberg, J., et al. Syndromic surveillance for bioterrorism following the attacks on the World Trade Center – New York City, 2001. Available from: URL: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm51Pa5.htm.
7. Farzad Mostashari, Annie Fine, Debjani Das et al. Use of ambulance dispatch data as an early warning system for community wide influenza like illness, New York City. Journal of Urban Health 2003; 80: Supplement 1: 43-49.
8. Micheal D, Pavlin JA, Mansfield LJ et al. Disease outbreak detection system using syndromic data in the greater Washington DC area. Am J Prev Med 2002; 23: 180-186.
9. Osaka, K, Takahashi H, Ohyama T. Testing a symptom based surveillance system at high profile gatherings as a preparatory measure for bio-terrorism. Epidemiol Infect 2002; 129: 429-434.
10. Willocks L, Crampin A, Milne L, et al. A large outbreak of cryptosporidiosis associated with a public water supply from a deep chalk borehole. Commun Dis Public Health 1998; 1: 239-243.
11. Armitage P, Berry G. Statistical method in medical research. 2nd ed. Oxford: Blackwell Scientific Productions, 1987.
12. Foods Standards Agency. A report of the study of infectious intestinal disease in England. Norwich: The Stationery Office, 2000.
13. Morbidity statistics from general practice: fourth national study, 1991-1992. London: HMSO, 1995.
14. The Birmingham Research Unit of the Royal College of General Practitioners. Annual Report 1999. Birmingham: RCGP, 1999.
15. Reis BY, Mandl KD. Time series modelling for syndromic surveillance. BMC Med Inform Decis Mak 2003; 23: 3.
16. Reis BY, Pagano M, Mandl KD. Using temporal context to improve biosurveillance. Proceedings of the National Academy of Sciences of the United States of America 2003; 100: 1961-1965.
17. Directorate of Access and Choice. Developing NHS Direct: A strategy document for the next 3 years. London: Department of Health, 2003.
18. Wagner MM, Data V, Dowling JN et al. Representative threat for research in public health surveillance. J Biomed Inform 2003; 36: 177-188
19. Johnson NL, Kotz S, Balakrishnan N. Continuous Univariate Distributions. 2nd ed. New York: Wiley, 1994: Chap18.

Annexe

Control chart method

The model for estimating the baseline incorporated a bank holiday factor, seasonal factor and day factor (weekday, Saturday or Sunday). The estimated baseline number of diarrhoea calls, as a proportion of total calls, is displayed in Figure 2. Using this estimated baseline we calculated a daily 99.5% upper prediction limit for the proportion of diarrhoea calls (Figure 2).

Confidence interval method

Within the NHS Direct syndromic surveillance system, there are NHS Direct sites and algorithms for which, due to resource constraints, 99.5% control chart limits are not calculated [3]. For these algorithms/sites, we calculate a seasonal 99.5% upper confidence interval, derived from standard formula for proportions [11]. The upper confidence interval was based on a seasonal baseline proportion of diarrhoea calls and remained constant throughout each season (i.e. not incorporating the bank holiday or day factor).

Simulation

The estimated baseline number of daily NHS Direct diarrhoea calls was generated from a Poisson distribution (θ) using the Gamma distribution (α, β) to account for over-dispersion (i.e. more variation than would be expected) [19].

α = estimated baseline / (scale – 1)

β = scale – 1

The scale parameter was calculated as [model deviance] / [model degrees of freedom], where the model was that used to calculate the baseline (expected value) from the control chart methodology. For the confidence interval method, the same scale parameter as derived from the control chart method was used. The statistical package STATA (version 8) was used for this work.

6 Can syndromic thresholds provide early warning of national influenza outbreaks?

Accepted as: Cooper DL, Verlander NQ, Elliot AJ, Joseph CA, Smith GE. Can syndromic thresholds provide early warning of national influenza outbreaks? *J Public Health (advance access published online on November 20, 2007).*

Abstract

Background: Influenza incidence thresholds are used to help predict the likely impact of influenza and inform health professionals and the public of current activity. We evaluate the potential of syndromic data (calls to a UK health helpline NHS Direct) to provide early warning of national influenza outbreaks.

Methods: Time series of NHS Direct calls concerning 'cold/flu' and fever syndromes for England and Wales were compared against influenza-like-illness (ILI) clinical incidence data and laboratory reports of influenza. *Poisson* regression models were used to derive NHS Direct thresholds. The early warning potential of thresholds was evaluated retrospectively for 2002-2006 and prospectively for winter 2006-2007.

Results: NHS Direct 'cold/flu' and fever calls generally rose and peaked at the same time as clinical and laboratory influenza data. We derived a national 'cold/flu' threshold of 1.2% of total calls and a fever (5-14 years) threshold of 9%. An initial lower fever threshold of 7.7% was discarded as it produced false alarms. Thresholds provided two weeks advanced warning of seasonal influenza activity during three of the four winters studied retrospectively, and 6 days advance warning during prospective evaluation.

Conclusions: Syndromic thresholds based on NHS Direct data provide advance warning of influenza circulating in the community. We recommend that age-group specific thresholds be developed for other clinical influenza surveillance systems in the UK and elsewhere.

Introduction

To help public health teams predict and respond to the likely impact of influenza the concept of thresholds levels have been applied in a number of national surveillance programmes [1-3]. Statistical threshold levels provide a clear and consistent picture to public health practitioners, the public, and the media about current influenza activity [4]. In England and Wales threshold levels have been estimated based on the incidence rate of new episodes of influenza-like illness (ILI) recorded by the Royal College of General Practitioners (RCGP) Weekly Returns Service (WRS). The RCGP WRS is a sentinel surveillance network currently comprised of 103 practices and approximately 500 general practitioners (GPs) monitoring an average weekly population of 900,000 patients [5]. A weekly ILI incidence rate of <30 per 100,000 indicates 'baseline' activity, 30-200 per 100,000 'normal seasonal' activity, and >200 per 100,000 'epidemic' activity.[1] These thresholds are used by the UK Department of Health to determine the decision on when GPs are advised to use antiviral drugs for treating seasonal influenza (e.g. oseltamivir) [6].

In recent years another source of clinical information has contributed to the UK influenza surveillance programme; syndromic calls to NHS Direct, a national nurse led telephone health line [7]. NHS Direct calls are broadly representative of the population although there is a high proportion of calls about young children, and higher use by women [8]. Previous work has shown that statistically significant rises in respiratory complaints to NHS Direct are timely indicators of influenza circulating in the community [9]. To better understand the value of syndromic data we compared NHS Direct data against clinical and laboratory influenza data over multiple influenza seasons. We then used a statistical model of NHS Direct calls and GP consultations about ILI to define NHS Direct threshold levels for influenza surveillance. These threshold levels were evaluated retrospectively and prospectively to test whether they provided advanced warning of community rises in influenza.

Methods

Data sources

NHS Direct nurses use clinical decision support software (the NHS Clinical Assessment System (NHS CAS)) to handle callers reported health problems. We obtained data for England and Wales, for weekly counts of NHS Direct calls classified under the 'cold/flu' clinical algorithm (all ages; 0-4yrs; 5-14yrs; 15-44yrs; 45-64yrs; 65+yrs) and fever algorithm (0-4yrs and 5-14yrs), and total algorithm calls, for weeks 40/02 to 20/06 inclusive. Weekly NHS Direct data were used, as opposed to daily data, to allow a direct comparison with weekly clinical and laboratory data. We obtained weekly ILI rates for England and Wales from the RCGP WRS for the same time period and age groups (from here on described as GP ILI). Weekly counts of the number (all ages combined) of positive influenza samples from community sources were collected from the Health Protection Agency, Centre for Infections virological surveillance scheme. This surveillance scheme collects five to 10 swabs per week from a network of 50 general practices in England and Wales. Data were compared using weekly time-series graphs.

Calculating NHS Direct thresholds

The relationship between weekly NHS Direct 'cold/flu' and fever calls, and GP ILI was explored using *Poisson* regression models, taking the total weekly number of 'cold/flu' or fever calls as the dependent variable, the logarithm of the total weekly number of calls as an offset, and GP ILI rate per 100,000 as the independent variable. Only weeks during the influenza season were used for the statistical models (weeks 40-20: October to May). *Poisson* regression is suitable for this analysis as the data of interest are events (NHS Direct calls) within specified time periods (weeks), and both numerator and denominator are reasonably high. The goodness of fit and validity of the models were tested using pseudo R^2 values, the Pearson chi squared statistic and inspection of model residuals.

A predicted NHS Direct upper threshold value for baseline activity (the weekly percentage of calls for cold/flu or fever calls) was calculated from the fitted model for a GP rate of 30 per 100,000 (the current clinical threshold for baseline activity). Separate models were constructed for each age-group and for cold/flu and fever calls separately. It is feasible that an NHS Direct threshold population rate, as well as a threshold percentage, may be useful for public health practice so separate models were also constructed using the logarithm of the total population of England and Wales as an offset. NHS Direct upper threshold rates per 100,000 for baseline activity were then calculated.

An NHS Direct threshold level for epidemic activity could not be estimated because there has not been 'epidemic' levels of GP ILI since winter 1999/2000 (before the NHS Direct national syndromic surveillance system was established).

Evaluating thresholds

The usefulness of NHS Direct thresholds (potential to provide early warning) was tested retrospectively for the study period (2002-2006), and prospectively during the winter of 2006/2007. For the retrospective evaluation we compared, for each age group, the week the derived NHS Direct threshold was exceeded (an NHS Direct alarm) against the week the existing GP ILI threshold of 30 per 100,000 was exceeded (GP alarm). NHS Direct alarms occurring in advance of any laboratory influenza detections were considered false alarms.

The derived NHS Direct threshold values were also used prospectively during winter 2006/2007. For the prospective evaluation we wanted to maximise the benefit of receiving and analysing daily NHS Direct data. Seven day rolling averages of the proportions of cold/flu and fever calls were calculated on a daily basis. An NHS Direct alarm was called if there were three consecutive daily alarms using the rolling averages. The date of the NHS Direct alarm was reported and was compared against the date that the first GP alarm was reported.

Results

During the study period there were five influenza seasons (Table I) of which 2003/2004 had the highest rates of GP ILI (peak = 62 per 100,000), and 2005/2006 the highest number of influenza reports (293 reports).

Comparison of data sources

The rise and peaks in the percentage of NHS Direct cold/flu calls and GP ILI occurred concurrently during winters 2003/2004 and 2004/2005 (influenza A dominant seasons) (Figure 1). There was a moderate peak in 'cold/flu' calls after relatively low peak in GP ILI during winter 2002/2003, and generally low levels of 'cold/flu' calls during winter 2005/2006. The timing of the rise and peaks in cold/flu calls for adults (15-64 yrs) was similar to total cold/flu calls (Figure 2). The lowest level of cold/flu calls was for the ≥65yr age-group mirroring the pattern of total NHS Direct usage for this age group [8]. For the 5-14yr age group, 'cold/flu' and fever calls peaked at the same time as laboratory reports of influenza, and within 1 week of the peak in GP ILI, during winters 2002/2003, 2003/2004, 2005/2006 (Figure 3). There were, however, also regular peaks of fever calls (5-14yrs) of between 8-9% during late December and July. 'Cold/flu' (0-4yrs) calls exhibited a secular downward trend and fever calls (0-4yrs) peaked at the same time as other influenza data during a single season in November 2003 (data not shown).

Fig 1. Weekly Royal College of General Practitioners (RCGP) weekly consultation rate for influenza-like-illness (ILI) per 100,000 (all ages), percentage of NHS Direct 'cold/flu' calls (all ages), number of influenza reports (CfI community scheme), and NHS Direct cold/flu threshold for baseline activity (2002-2006).

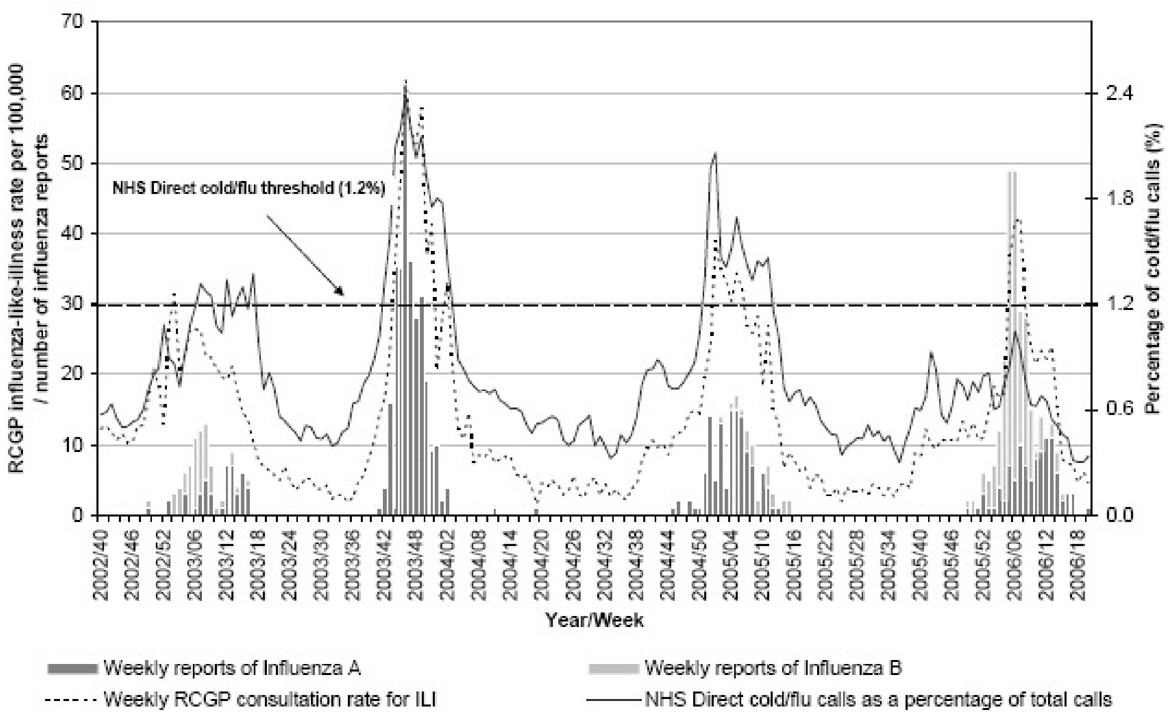
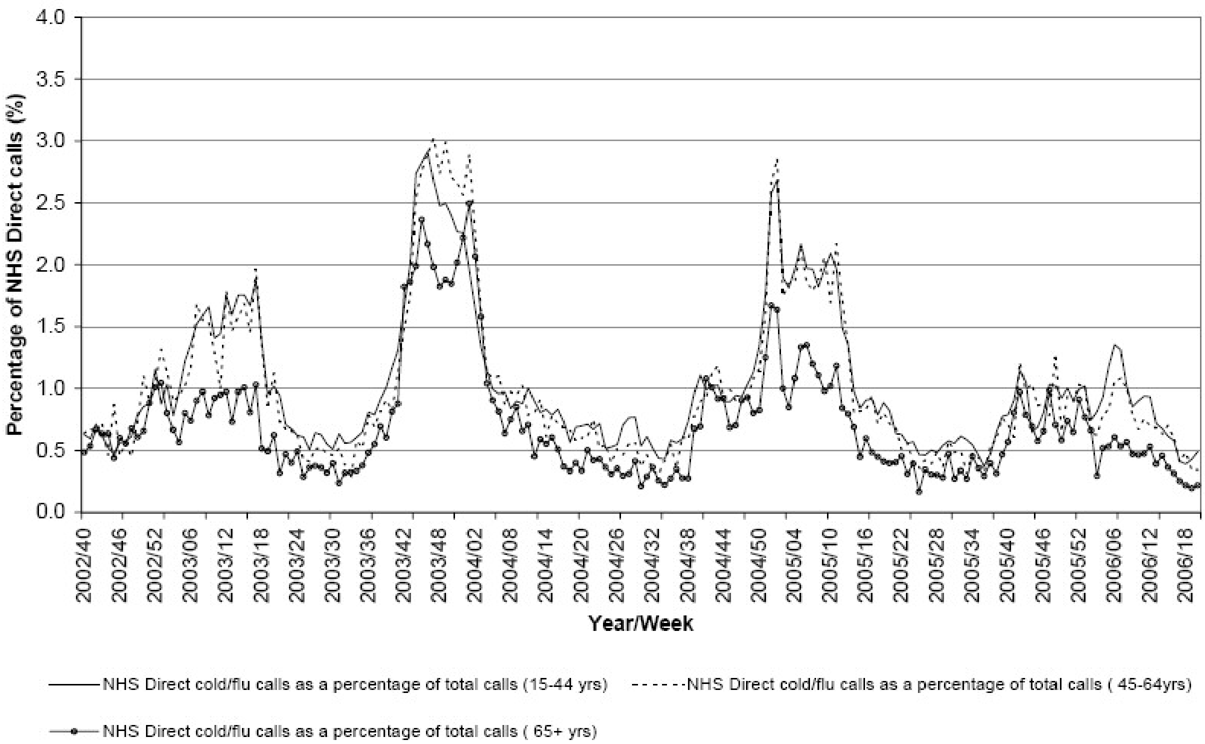


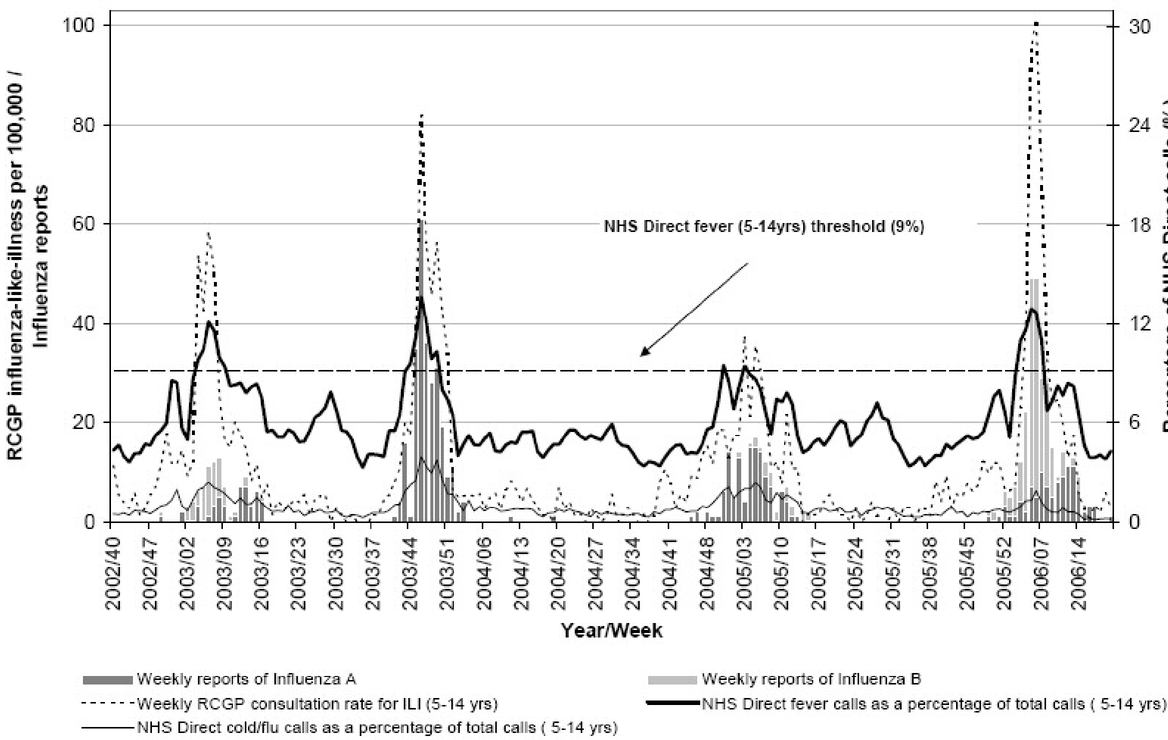
Fig 2. Weekly percentage of NHS Direct 'cold/flu' calls by adult age-group (2002-2006).



Threshold values

The best fitting *Poisson* models were for the percentage of 'cold/flu' calls (all ages) (pseudo R^2 =0.54), rate of cold/flu calls (all ages) (R^2=0.51), proportion of fever calls (5-14yrs) (R^2 = 0.53), and rate of fever calls (5-14yrs) (R^2 = 0.51) (Table II). With a *Poisson* model the pseudo R^2 values provide an approximate measure of the explanatory power of the model, but not an exact measure of the proportion of variation explained by the model, as for linear regression models. High Pearson χ^2 statistics for these same models revealed that the data did not conform to a standard *Poisson* distribution and was likely to be over-dispersed as previously demonstrated [8]. Examination of the model residuals revealed that the cold/flu model over estimated the percentage of calls during winter 2005/2006, and that the fever (5-14yr) model under estimated the percentage of calls during weeks 50 to 52, whilst over estimating during non-influenza periods (Figure 4).

Fig 3. Weekly RCGP ILI consultations per 100,000 (5-14 yrs), percentage of NHS Direct 'cold/flu' (5-14yrs) and fever calls (5-14yrs), number of influenza reports (CfI community scheme), and NHS Direct thresholds for fever (5-14ys) calls (2002-2006).



Through a combination of visually examining trends in the three data sources and using the best fitting *Poisson* models, three indicators - cold/flu calls (all ages), cold/flu calls (5-14yrs) and fever calls (5-14yr)) - appeared most suitable for testing the early warning potential of NHS Direct thresholds. These three indicators are now examined in more detail.

The models estimated an NHS Direct 'cold/flu' upper threshold for baseline activity of 1.2% of total calls or 1.5 calls per 100,000 (Table II, Figure 1), comparable to the GP ILI threshold of 30 per 100,000. The estimated cold/flu (5-14yrs) threshold was 1.3% or 1.25 per 100,000, and fever (5-14yrs) 7.7% or 7.6 per 100,000 (Table II, Figure 3). Due to the regular December peaks in the proportion of fever calls (5-14yrs), in excess of the 7.6% threshold, we decided to raise this threshold to 9% to avoid false alarms during periods when laboratory data indicated that influenza was not circulating. These numerical threshold values represent the transmission point from baseline to normal seasonal influenza activity.

Table I - Summary of influenza surveillance data from laboratories, general practitioners (GPs), and NHS Direct during the study period (2002-2007).

	Laboratory: Virological surveillance scheme				Clinical: GP influenza-like-illness (ILI)		NHS Direct cold/flu calls	
Winter	Dominant influenza subtype	Total number of positive reports	Peak week	Peak week number of reports	Peak week	Peak week rate per 100,000	Peak week	Peak week percentage of calls (%)
2002/2003	A (H3N2) and B co-circulation	101	2003/08	13	2003/02	31	2003/17	1.5
2003/2004	A(H3N2)	258	2003/46	61	2003/46	62	2004/46	2.4
2004/2005	A(H3N2)	143	2005/05	17	2005/01	38	2005/01	2.1
2005/2006	B and A(H3N2) co- circulation	293	2006/05& 06	49	2006/07	44	2006/06	1
2006/2007	A(H3N2)	*			07/2007	44	2007/06	1.1

***** Data not yet available for the entire influenza season***

Table II. Existing upper threshold value for baseline activity of GP consultations for ILI, newly derived NHS Direct threshold values for cold/flu and fever calls derived, model pseudo R^2 values.

	Indicator	Threshold value for baseline activity	Pseudo R-squared (R^2)
GP consultations	GP ILI (all ages)	30 per 100,000	-
NHS Direct cold/flu calls	NHS Direct cold/flu (all ages)	1.2% of total calls	0.54
		1.5 calls per 100,000	0.51
	NHS Direct cold/flu (0-4yrs)	0.2% of total calls	0.20
		1 per 100,000	0.25
	NHS Direct cold/flu (5-14yrs)	1.3% of total calls	0.31
		1.25 per calls 100,000	0.43
	NHS Direct cold/flu (15-44yrs)	1.4% of total calls	0.48
		2 per 100,000	0.42
	NHS Direct cold/flu (45-64yrs)	1.6%	0.44
		1.2 per 100,000	0.43
	NHS Direct cold/flu (≥65yrs)	1.6% of total calls	0.39
		1.1 per 100,000	0.30
NHS Direct fever calls	NHS Direct fever (0-4yrs)	1.2%	0.44
		57.9 per 100,000	0.28
	NHS Direct fever (5-14yrs)	7.7% of total calls	0.53
		7.6 per calls 100,000	0.51
	Raised NHS Direct fever (5-14yrs) threshold	9% of total calls	-
		9.5 calls per 100,000	-

Evaluation

Retrospective – winters 2002 to 2006

Table III presents, for each influenza season, the time lag between the first exceedance of the NHS Direct threshold for baseline activity (termed an NHS Direct alarm) and the first exceedance of the GP ILI threshold for baseline activity (termed a GP alarm). During winter 2002/2003 the NHS Direct cold/flu (all ages) alarm occurred four weeks after the GP alarm. The NHS Direct 'cold/flu' (5-14yrs) alarm occurred three weeks prior to the GP alarm. This is considered a false alarm as laboratory data indicated no influenza activity during this week (Figure 3). During winters 2003/2004 and 2004/2005 all 3 NHS Direct alarms occurred one to three weeks in advance of the GP alarm. During winter 2005/2006 the raised NHS Direct fever (5-14yr) alarm occurred two weeks in advance of the GP alarm.

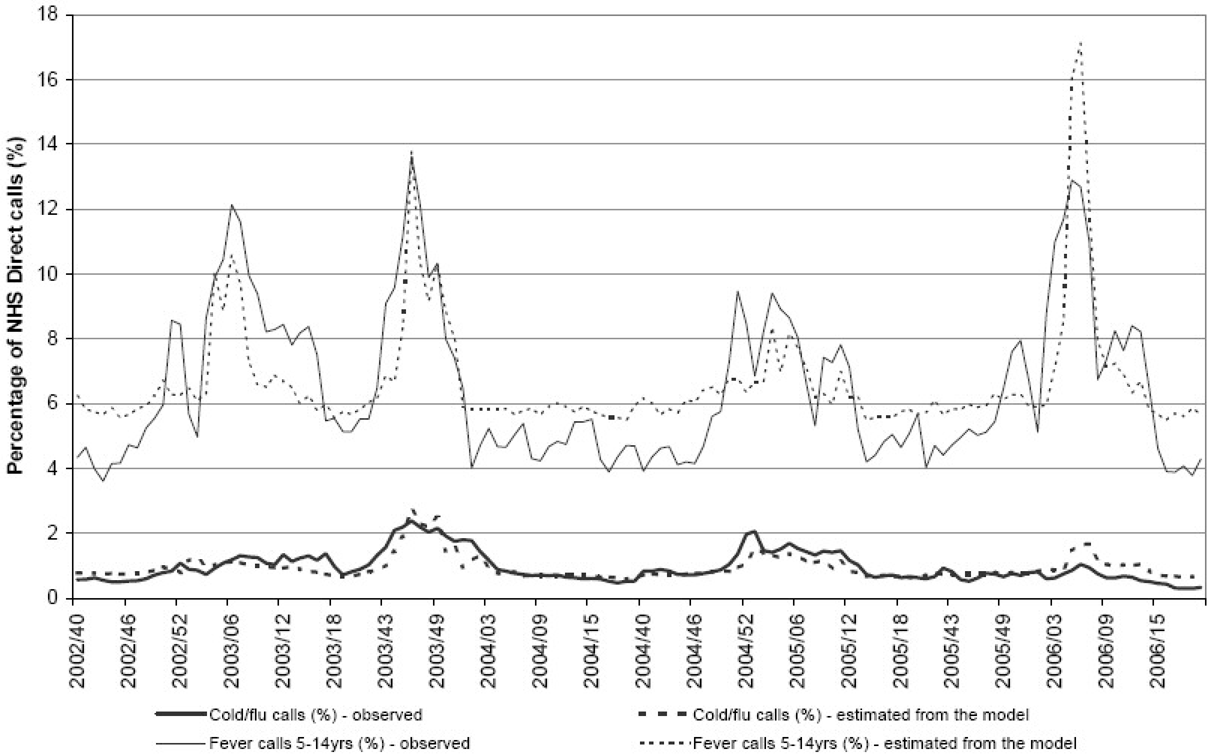
Table III. Time lag (weeks) between the RCGP threshold being exceeded (GP alarm) and the NHS Direct threshold being exceeded (NHS Direct alarm) for the retrospective evaluation. For example, a time lag of -2 indicates that the NHS Direct alarm occurred 2 weeks in advance of the GP alarm.

		Influenza season			
	Winter	2002/2003	2003/2004	2004/2005	2005/2006
	Week of the first GP ILI alarm	Year/Week 2003/02	Year/ Week 2003/44	Year/Week 2005/01	Year/Week 2006/05
Time lag between an NHS Direct alarm and a GP ILI alarm	NHS Direct cold/flu (all ages)	+4	-2	-2	No alarm
	NHS Direct cold/flu (5-14yrs)	-3	-1	-3	+1
	Raised NHS Direct fever (5-14yrs) threshold	+2	-1	-2	-2

Prospective – winter 2006/2007

The first NHS Direct fever (5-14yr) alarm occurred on the 31st January 2007, was reported to national influenza surveillance coordinators on the 1st February, and in a publicly available surveillance bulletin. The GP weekly alarm signalled during week 05/01 (30.2 consultations of ILI per 100,000) and was reported on 7th February. The NHS Direct cold/flu (all ages) and cold/flu (5-14yrs) alarm did not signal during the winter of 2006/2007, and the cold/flu (15-44 yrs) alarm signalled on the 9th of February.

Fig 4. Observed and fitted weekly values for the cold/flu calls (all ages) and fever calls (5-14yrs) *Poisson* models.



Discussion

Main findings

NHS Direct calls about 'influenza-like-illness' generally rose and peaked at the same time as GP diagnosed ILI and laboratory reports of influenza. Influenza thresholds based on NHS Direct calls provided advance warning of influenza circulating in the community. We propose an overall 'cold/flu' threshold of 1.2% of total calls and fever threshold of 9% for the 5-14 year age-group be used for influenza surveillance in England and Wales. In a retrospective evaluation these two thresholds provided two weeks advanced warning of seasonal influenza activity in three of four winters studied, with no false alarms. During prospective evaluation the NHS Direct fever threshold signalled six days earlier than an influenza surveillance system based on GP consultations. Timely information was given to influenza surveillance coordinators, local health protection teams and the public.

The use of age-group specific thresholds meant that the start of the influenza season was detected regardless of the influenza subtype and age groups predominantly affected. No single age-group threshold provided early warning during all winters. The threshold value for total 'cold/flu' calls was of particular value during a winter characterised by influenza activity early in the 2003/2004 season of a new influenza A drift variant (A/Fujian/411/2002 (H3N2)-like virus [10]). In addition, the threshold value for fever calls (5-14 year) provided a comparative advantage during a national influenza B outbreak during 2005/2006. Variation in the amount of early warning provided by these thresholds (ranging from 0 to 2 weeks in our study) is likely to have been influenced by annual changes in the virulence of the circulating influenza sub-types and the main age-group affected, which in turn influence presenting patterns to NHS Direct and GP surgeries. We suggest that a range of age-group specific thresholds are required for timely warning of different subtypes of the influenza virus, and are recommended for surveillance systems based on telehealth data.

What is already known on this topic

Influenza-like-illness is caused by infection or co-infection with a number of pathogens including influenza, respiratory syncytial virus (RSV), adenovirus and parainfluenza. In the UK, RSV regularly peaks in late December so confounds these results. For example, 'NHS Direct fever peaks' occurred during week 50 in all but the 2003/2004 winter when it was submerged within the early peak caused by the new influenza drift variant. It is likely the confounding effect of RSV and other pathogens is greater for NHS Direct calls than GP ILI as the NHS Direct baseline level is relatively higher during non-influenza periods (Figure 1). Also NHS Direct nurses triage callers using a generic 'cold/flu' algorithm so specificity is likely to be lower than GP ILI where RSV infection is predominately diagnosed as acute bronchitis [11]. Previous modelling work suggests RSV, influenza and pneumococcal are the main cause of respiratory complaints made to NHS Direct [12]. A self-sampling study of NHS Direct callers previously resulted in 16% testing positive for influenza and 5% for RSV during winter [13]. Despite the apparent 'NHS Direct RSV peak' these results suggest that the NHS Direct thresholds have sufficient specificity to minimise false alarms caused by circulation of other respiratory diseases.

Numerical influenza thresholds are used widely throughout Europe [14] and in the US [15] to describe influenza activity. They provide a clear and consistent message of influenza levels, place current activity in the contact of previous winters and avoid confusion caused by different authorities describing influenza levels in different ways. Thresholds are also used to trigger action by GPs to prescribe antiviral medication, aid diagnoses, and encourage identification of high risk patients [4] and sampling for laboratory testing. Communications may also be made with hospitals forewarning them of the need to free bed space for an expected rise in respiratory admissions. The necessity for accurate influenza thresholds is such that the Health Protection Agency is currently recommending that a variety of surveillance systems (including NHS Direct) should be used to trigger antiviral prescribing for influenza, after the realisation that virus activity can occur before current thresholds levels are reached and after they have fallen back to baseline [1].

Various means of obtaining statistical thresholds such as ARIMA regression models [16] and CUSUMS [17] have been employed for syndromic surveillance purposes. These methods commonly use Emergency Department data and detect significant changes in daily visits from an expected level. Often initially designed for the detection of potential bio-terrorist events [18], these systems have been successful in detecting statistically significant rises in respiratory disease and the onset of the influenza season [19].

What this study adds

The syndromic thresholds derived here differ from other methods in that they are based on telehealth data in comparison with influenza data from other sources (clinical and laboratory). They are also designed solely for detecting the *start* of influenza seasons and not for surveillance of general respiratory disease. Work in the US comparing telephone health line data against doctor diagnosed ILI and influenza isolates has shown concurrent peaks in data [20]. In another study, numbers of medical centre outpatients with ‘fever’ correlated strongly with viral respiratory pathogens [21]. Our work uses a different source of syndromic data (telehealth calls) but supports these conclusions, adding that a community rise in influenza may be predicted by a rise in fever calls. It is hoped that this work will be of relevance to evolving syndromic surveillance systems using telehealth data, for example in Canada

[22], and within the newly established framework for ‘epidemic intelligence’ gathering in Europe [23].

Although the RCGP GP surveillance scheme, used as a reference point, represents less than 2% of the population of England and Wales we consider this the gold standard for GP ILI surveillance having reported ILI rates since the 1960s [24]. We have shown age-group specific NHS Direct thresholds may complement this scheme. The fever calls threshold for school age children may be of particular use when the predominant circulating influenza strain does not initially cause significant levels of illness reported to general practitioners. For example, during January and February 2006 there were an unusually high number of school outbreaks (many confirmed as influenza B) in England and Wales, not initially reflected within the GP ILI rates [25]. It may be argued that the NHS Direct fever threshold, which signalled two weeks earlier than GP ILI during winter 2005/2006, was more sensitive in detecting this national influenza outbreak. In practice, NHS Direct data can be analysed and reported on each weekday. This extends the lead time as weekly GP ILI data are generally reported 3 days after the end of the reporting week, and laboratory data from community virological schemes up to 10 days from specimen collection. Further work is needed to test whether prospective spatio-temporal analyses of NHS Direct data to a local level could identify the origin of national outbreaks, thus extending the lead time further.

Limitations of this study

Systematic residual patterns suggest that a constant threshold value over an entire winter may not be justified and that an iterative approach, with daily or weekly changes in the threshold, is more suitable (e.g. ARIMA, CUSUMS). The NHS Direct surveillance system already employs a control chart methodology incorporating day of the week, seasonal and long term trend factors [9]. These control charts flag significant rises in a variety of syndromes (e.g cold/flu, fever, diarrhoea, vomiting). This study, however, has taken a pragmatic approach to produce constant threshold values designed specifically for influenza surveillance, derived directly from *Poisson* models and adjusted if necessary to improve specificity. It is hoped that these thresholds will complement the control charts already employed. Similarly, pragmatism has also driven derivation of the RCGP threshold for baseline ILI activity which was recently lowered from <50 to <30 per 100,000 following studies of 10 years incidence data [1,26]. By examining the model residuals along with the time lag between signals (Table III) we were able to determine which age-group specific threshold was most useful in differing influenza seasons (e.g. influenza A or B).

Relatively low levels of influenza activity in the UK during the study period has made it difficult to derive thresholds for epidemic activity. For example during winter 1999/2000 GP ILI peaked at 256 per 100,000, four times higher than the peak rate of 62 per 100,000 during our study period. A principle aim of this paper, however, was to establish thresholds for early warning (notification of the beginning of the influenza season) and not to predict extreme values. The secular downward trend in respiratory illness reported in the UK [1] means there may not be the opportunity to test our syndromic surveillance system in epidemic conditions until novel influenza viruses (e.g. avian influenza A(H5N1)) with high human pathogenicity emerge.

References

1. Goddard NL, Kyncl J, Watson JM. Appropriateness of thresholds currently used to describe influenza activity in England. *Commun Dis Public Health* 2003 Sep; 6(3):238-45.
2. Quenel P, Dab W. Influenza A and B epidemic criteria based on time-series analysis of health services surveillance data. *Eur J Epidemiol* 1998 Apr; 14(3):275-85.
3. Watts CG, Andrews RM, Druce JD, Kelly HA. Establishing thresholds for influenza surveillance in Victoria. *Aust N Z J Public Health* 2003; 27(4):409-12.
4. Dedman DJ, Watson JM. The use of influenza thresholds to describe levels of influenza activity. *PHLS Microbiology Digest* 1997; 14(4): 206-208.
5. Fleming DM. Weekly Returns Service of the Royal College of General Practitioners. *Commun Dis Public Health* 1999 Jun; 2(2):96-100. Review.
6. Salisbury D. Influenza season 2005/06 – use of Oseltamivir. February 2006; Department of Health.
7. NHS Direct Online website. http://www.nhsdirect.nhs.uk/. Accessed April 2007.
8. Payne F, Jessopp L. NHS Direct: review of activity data for the first year of operation at one site. *J Publ Hlth Med* 2001; 23: 155-158.
9. Baker M, Smith GE, Cooper DL, et al. Early Warning and NHS Direct: a role in community surveillance? *J Publ Hlth Med* 2004; 25:362-368.
10. Cooke MK, Crofts JP, Joseph CA, et al. Influenza and other respiratory viruses survellance in the United Kingdom October 2003 to May 2004. *Commun Dis Rep Suppl* 2005; 15(1): 1-8.
11. Fleming DM, Elliot AJ, Cross KC. Morbidity profiles of patients consulting during influenza and respiratory syncytial virus active periods. *Epidemiol Infect* 2007; doi:10.1017/S0950268807007881.
12. Cooper DL, Smith GE, Edmunds WJ, Joseph C, George RC. The contribution of respiratory pathogens to the seasonality of NHS Direct calls. J Infect (in press).
13. Cooper DL, Smith GE, Chinemana F, et al. Linking syndromic surveillance with virological sampling. Epidemiol Infect (in press).
14. Aguilera JF, Paget J, Manuguerra JC. Survey of Influenza Surveillance Systems in Europe. European Influenza Surveillance Scheme Website 2001. http://www.eiss.org/documents/eiss_inventory_survey_dec_2000.pdf
15. CDC. Influenza Activity, United States and Worldwide, 2005-06 Season, and Composition of the 2006-07 Influenza Vaccine. *MMWR* 2006; 55(23):648-653.
16. Reis BY, Mandl KD. Time series modelling for syndromic surveillance. *BMC Med Inform Decis Mak* 2003 Jan 23; 3:2.
17. Hutwagner LC, Thompson WW, Seeman GM, Treadwell T. A simulation model for assessing aberration detection methods used in public health surveillance for systems with limited baselines. *Stat Med* 2005 Feb 28; 24(4):543-50.
18. Buehler JW, Berkelman RL, Hartley DM, Peters CJ. Syndromic surveillance and bioterrorism-related epidemics. *Emerg Infect Dis* 2003 Oct; 9(10):1197-204.
19. Heffernan R, Mostashari F, Das D, Karpati A, Kulldorff M, Weiss D. Syndromic surveillance in public health practice, New York City. *Emerg Infect Dis* 2004; 10(5): 858-864.
20. Yih K, Teates KS, Abrams A, Ken Kleinman, et al. Nurse Call Data for Detection of Influenza-Like Illness. *Advances in Disease Surveillance* 2006; 1:76.
21. Marsden-Haug N, Foster VB, Gould PL, et al. Code-based Syndromic Surveillance for Influenzalike Illness by International Classification of Diseases, Ninth Revision. *Emerg Infect Dis* 2007; 13(2): 207-216.
22. Rolland E, Moore KM, Robinson VA, McGuinness D. Using Ontario's "Telehealth" health telephone helpline as an early-warning system: a study protocol. *BMC Health Serv Res* 2006;6:10.
23. Paquet C, Coulombier D, Kaiser R, Ciotti M. Epidemic intelligence: a new framework for strengthening disease surveillance in Europe. *Euro Surveill* 2006; 11(12):E061221.3.
24. Elliot AJ, Fleming DM. Surveillance of influenza-like illness in England and Wales during 1966-2006. *Euro Surveill* 2006; 11(10)249-50.
25. Desai S, Zhao H, Cooke MK, et al. Surveillance of Influenza and other respiratory viruses in the United Kingdom: October 2005 to May 2006. *Commun Dis Rep Suppl* 2006; 16(39): 1-6.
26. Fleming DM, Zambon M, Bartelds AIM, de Jong JC. The duration and magnitude of influenza epidemics: a study of surveillance data from sentinel general practices in England, Wales and the Netherlands. *Euro J Epidemiol* 1999; 15(5):467-73.

7 The contribution of respiratory pathogens to the seasonality of NHS Direct calls

Published as: Cooper DL, Smith GE, Edmunds WJ, Joseph C, Gerard E, George RC. The contribution of respiratory pathogens to the seasonality of NHS Direct calls. *J Infect 2007; 55:240-248.*

Abstract

Objectives: Primary care is thought to bear half the cost of treating infections in the UK. We describe the seasonal variation in NHS Direct respiratory calls (a new source of primary care data) and estimate the contribution of specific respiratory pathogens to this variation.

Methods: Linear regression models were used to estimate the weekly contribution of specific respiratory pathogens to the volume of NHS Direct respiratory calls (England and Wales, 2002-2004, all ages and 0-4 years).

Results: Annual peaks in NHS Direct cough and difficulty breathing calls occurred in late December, with peaks in 'cold/flu' and fever calls occurring between November and April. The main explanatory variables were influenza (estimated to account for 72.5 calls per 100,000/year; 22% of 'cold/flu' calls; 15% of cough; and 13% of fever) and Streptococcus pneumoniae (55.5 per 100,000; 33% of 'cold/flu' calls; 20% of cough; and 15% of fever (0-4 years)).

Conclusions: It is estimated that respiratory viruses, notably influenza and RSV, are responsible for at least 50% of the seasonal variation in NHS Direct respiratory calls. These results provide estimates of the burden of specific respiratory diseases reported to NHS Direct, and will help interpret syndromic surveillance data used to provide early warning of rises in community morbidity.

Introduction

Primary care is thought to bear over half the cost of treating infections in the UK (£6 billion per year) with 5.5 million GP consultations for respiratory infections alone [1]. NHS Direct [2] - a nurse led national telephone health helpline established in 1999 – provides the opportunity to examine a new source of reported respiratory disease in England and Wales, some of which may go un-recorded by traditional information and surveillance systems (GP and hospital). Published work about NHS Direct 'respiratory' calls' suggests that some of the temporal variation may be related to the seasonality of influenza [3]. Many other pathogens (viral and bacterial) may cause respiratory disease, however, and the relative contribution of these pathogens to NHS Direct calls has not been estimated. Describing and quantifying the burden of respiratory disease is necessary to identify health priorities and provide evidence to support future research and development.

Annual rises in respiratory diseases such as influenza, the emergence of new infections such as SARS and 'avian flu', and the potential of bio-terrorist attacks to cause widespread disease, means timely and representative disease surveillance data are desirable. NHS Direct call centre data have been monitored for surveillance purposes since 2001 – the NHS Direct Syndromic Surveillance System [4] – in order to provide early warning of a rise in community morbidity. Syndromic surveillance systems use a variety of pre-diagnostic data sources (e.g. ambulance encounters [5], A&E admissions [6], pharmacy sales [7], calls to health help-lines [4] in an attempt to identify outbreaks and incidents of public health importance. Although NHS Direct syndromic data have alerted public health teams to national or local rises in illness [4] there is no mechanism for routine microbiological confirmation of the cause of NHS Direct calls about presumed infections. Comparison of trends in NHS Direct calls against laboratory reports may, therefore, help elucidate a likely cause of NHS Direct respiratory calls, and help interpret surveillance trends.

In this report we describe the seasonal variation in NHS Direct respiratory calls and use a multiple linear regression model to estimate the contribution of different respiratory pathogens to this variation. A similar technique has been used previously to estimate the contribution of respiratory syncytial virus (RSV) and other respiratory pathogens to hospital admissions for lower respiratory tract infections [8], and the contribution of pneumococcal disease to GP consultations [9].

Methods

Weekly counts of NHS Direct calls about 'cold/flu' (CF), cough (C), fever (F), and difficulty breathing (DB) in England and Wales were extracted from the NHS Direct syndromic surveillance database for the period October 2002 to October 2004 (all ages and 0-4 years separately). The appropriateness of classifying fever as a respiratory call is discussed later. NHS Direct data for cough calls (0-4 yrs) were only available for December 2002 to October 2004 and difficulty breathing calls (0-4yrs) for April 2003 to October 2004. There are no formal case definitions for NHS Direct respiratory syndromes (as symptoms are self-reported and triaged remotely) and data extraction was based on the clinical algorithms ('cold/flu', 'cough', 'fever' or 'difficulty breathing') selected by the NHS Direct nurses handling the calls.

Table 1 - Numbers of NHS Direct calls for respiratory symptoms/syndromes in England and Wales (October 2002 – October 2004).

NHS Direct symptom/syndrome	Number of calls (all ages)	Number of calls (0-4yrs)
Fever	272,812	182,972
Cough	190,578	83,476 *
Difficulty breathing	77,317	22,471 **
Cold/flu	60,747	3,948
Total	**601,454**	**292,867**

* Data for cough calls (0-4yrs) weeks 52/02 to 45/04.

** Data for difficulty breathing calls (0-4yrs) weeks 15/03 to 45/04.

Weekly counts of laboratory reports (by earliest specimen date) of the following main causes of respiratory disease were extracted from the Health Protection Agency Centre for Infections database of laboratory reports for England and Wales (LabBase), for October 2002 to October 2004 (all ages and 0-4 years separately):

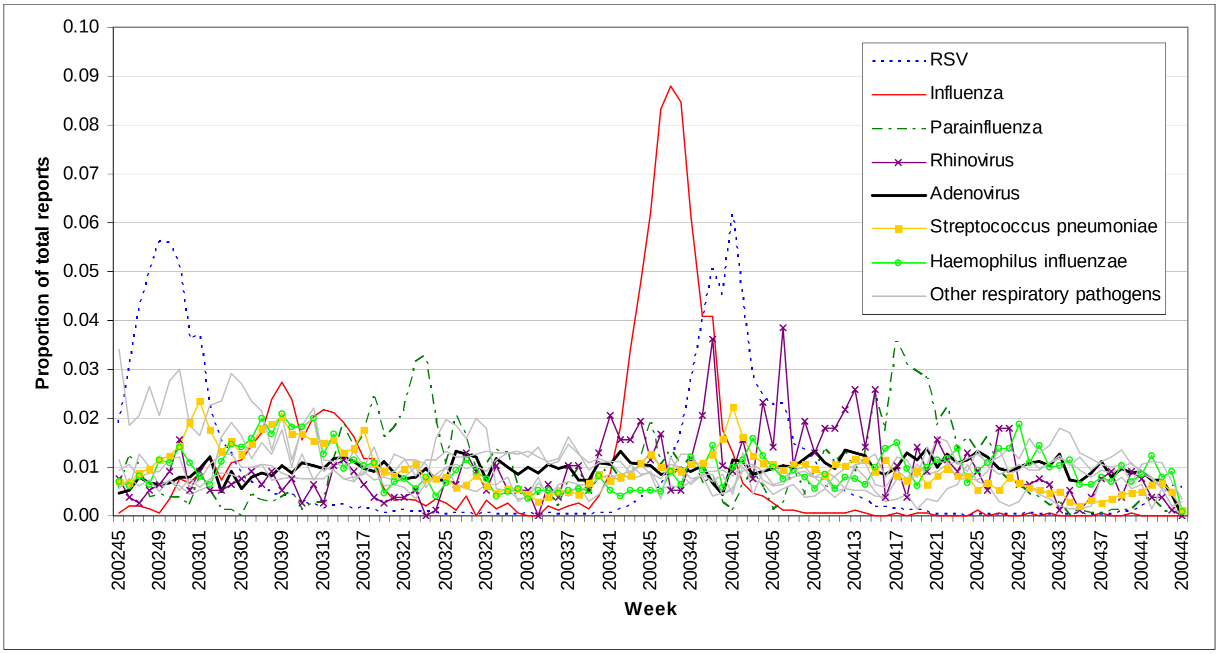
Respiratory syncytial virus (RSV)	Influenza
Parainfluenza	Rhinovirus
Adenovirus	Coronavirus
Invasive *Streptococcus pneumoniae*	*Mycoplasma pneumoniae*
Klebsiella pneumoniae	C*oxiella burnetii*
Legionella species	*Bordetella pertussis*
Invasive *Haemophilus influenzae*	C*hlamydia pneumoniae*
Chlamydia psittaci	*Moraxella catarrhalis*

LabBase is a national HPA Centre for Infections database that collects laboratory reports of all micro-organisms isolated at approximately 400 NHS and other laboratories throughout England, Wales, and Northern Ireland. Although enhanced and separate community surveillance systems exist for some pathogens (e.g. influenza [10], legionella [11]) a single source of laboratory data was preferred to reduce selection bias between different systems.

For each pathogen we compared trends in laboratory reports for all ages and 0-4 years. If both age groups exhibited a similar seasonality then all ages data were used to make the seasonal pattern more distinct. Two binary dummy variables – ‘Christmas bank holiday week’ and ‘other public holiday weeks’ – were created to account for excess out-of-hours demand for NHS Direct during public holiday periods (i.e. when many GP surgeries are closed).

Using methods previously described [12], multiple linear regression models were constructed to estimate the weekly proportion of NHS Direct respiratory calls (all ages and 0-4 years separately) attributable to respiratory pathogens (explanatory variables) (SPSS v13.0 was used). This method uses a regression model to estimate how the seasonal variation in laboratory reports is reflected within the variation in NHS Direct calls. Linear regression models assume that the relationship between the dependent and explanatory variables is linear and that the model errors are distributed normally and do not exhibit serial correlation. A constant number of NHS Direct calls due to other infectious or non-infectious disease (other causes) was assumed within each model. Variables that contributed little to the model were removed by backward stepwise regression. This technique gradually removes the variable that causes the smallest reduction in R^2 (goodness of fit of the model) until removal of a further variable causes a significant reduction in R^2 (by the F- test). Once non-significant variables were removed, any remaining significant variables with negative coefficients were also removed from the model as it was considered biologically implausible that they negatively influence the number of NHS Direct calls. The model was then re-estimated with the remaining variables (Model1). Possible model misspecification was tested by constructing alternative models to examine the changes in coefficients of the final model (Model1) to sequential deletion of significant variables. Model residuals were examined for autocorrelation.

Figure 1: Weekly variation in laboratory reports of respiratory pathogens (October 2002 – October 2004), displayed as a proportion of total reports for each pathogen. [Pathogens that significantly influenced the models are presented separately; pathogens with < 100 reports per year are not displayed].



Estimates were made of the proportion of NHS Direct calls attributable to each significant respiratory pathogen over the two year period and by week. We also estimated the proportion of NHS Direct calls attributable to influenza and RSV during periods of high activity (influenza and RSV 'seasons'). Influenza seasons were defined as continuous periods during which the weekly number of influenza reports exceeded 10 (this accounted for 92% of total influenza reports). RSV seasons were continuous periods when the weekly number of RSV reports exceeded 100 (87% of total RSV reports). From the estimated proportion of calls about specific pathogens, the estimated incidence per 100,000 of NHS Direct calls due to specific pathogens was calculated for England and Wales.

Results

NHS Direct respiratory calls

Between October 2002 and October 2004 there were 601,454 calls about the four respiratory syndromes, 45% classified as fever, 32% cough, 13% difficulty breathing and 10% ‘cold/flu (Table 1). These four syndromes comprised 8% of total NHS Direct calls. Of the 292,867 respiratory calls about children under 5 years, 62% were classified as fever, 29% cough, 8% difficulty breathing and 1% cold/flu. There were two peaks in calls about cold/flu and fever occurring in Feb-April 2003 and November 2003. Peaks in cough and difficulty breathing calls occurred at the end of December (week 52) during 2002 and 2003.

Laboratory reports of respiratory pathogens

For the two year period there were 49,652 laboratory reports, 20,257 (41%) of which were about children under 5 years. RSV was the most commonly recorded pathogen for all ages (13,544 reports) and children under 5 years (8,048). The seasonality of respiratory pathogens varied between those with distinct winter peaks (e.g. RSV): weekly range 1 to 835 reports) (Figure 1) and those with no clear seasonality (e.g. adenovirus: range 22 to 70 reports).

Model results

The models explained 79%, 91%, 71% of the seasonal variation in cold/flu (Model CF1), cough (Model C1), fever (Model F1) (Table 2). For the 0-4 year age group, the models explained 88% of the variation in cough calls (Model C1(0-4)) and 76% of the variation in Fever calls (Model F1(0-4)) (Table 2). Model CF1 (0-4 yrs) explained 75% of the variation on cold/flu calls. However, the NHS Direct cold/flu clinical algorithm is rarely used to handle calls about children under 5 years (averaging only

37 calls per week in England and Wales) so full results of Model CF1 (0-4yrs) are not presented in Table 2. Detailed results of the difficulty breathing models are also not reported as the power of both models was low (Model DB1: $R^{2} =$ 0.23, Model DB(0-4): $R^2 = 0.48$).

Estimated contribution of respiratory pathogens to NHS Direct calls

Table 2 presents the estimated proportion of NHS Direct calls attributable to specific respiratory pathogens. The best-fitting parsimonious model suggested that over the two year period, influenza accounted for 22% of cold/flu calls (all ages), 15% of cough calls (all ages), 19% of cough calls (0-4 years), 13% of fever calls (all ages) and 13% of fever calls (0-4 years), with RSV accounting for 12% and 16% of cough calls (all ages and 0-4 years). The contribution of influenza and RSV to NHS Direct respiratory calls, however, was confined mainly to the influenza season (weeks 51/02–19/03 and 41/03–02/04) and RSV seasons (45/02-05/03 and 47/03-09/04) (Figures 2-4). During the two influenza seasons influenza was estimated to account for 25% and 51% of cold/flu calls (all ages), and 17% and 41% of fever calls (0-4 years). RSV accounted for 19% and 37% of cough calls (all ages) and 25% and 38% of cough (0-4 years) during successive RSV seasons. The other main contributory variable was *S. pneumoniae*, estimated to account for 33% of cold/flu calls (all ages), 19% of fever calls (0-4yrs), and 20% of cough calls (all ages) over 2 years.

Figure 2-4 show the weekly estimates of the contribution of respiratory disease pathogens to NHS Direct cold/flu calls, cough calls (all calls and 0-4 years separately), and fever calls (0-4 years). They illustrate the relatively high but short lived contribution of influenza and RSV during seasonal bursts, and the year round contribution, with December/January peaks, of *S. pneumoniae*. The model estimates generally corresponded to the observed number of NHS Direct calls, although the cold/flu and cough models provided poor estimates for weeks 39/04 to 45/04 (Figures 2, 3a and 3b).

There appeared to be little confounding between diseases with similar seasonality and clinical presentation (e.g. influenza and RSV) as removing either of these variables from the final models resulted in only minor increases in the estimated contribution of the remaining variables. For example, dropping RSV, rhinovirus and parainfluenza from model C1 (0-4) led to a small estimated increase in the contribution of influenza from 19% to 23% (model C3(0-4)) (Table 2).

Figure 2. Comparison of the observed number of NHS Direct 'cold/flu' calls (all ages) with the estimated numbers due to influenza, *Streptococcus pneumoniae*, rhinovirus and other causes (October 2002 – October 2004), (model CF1).

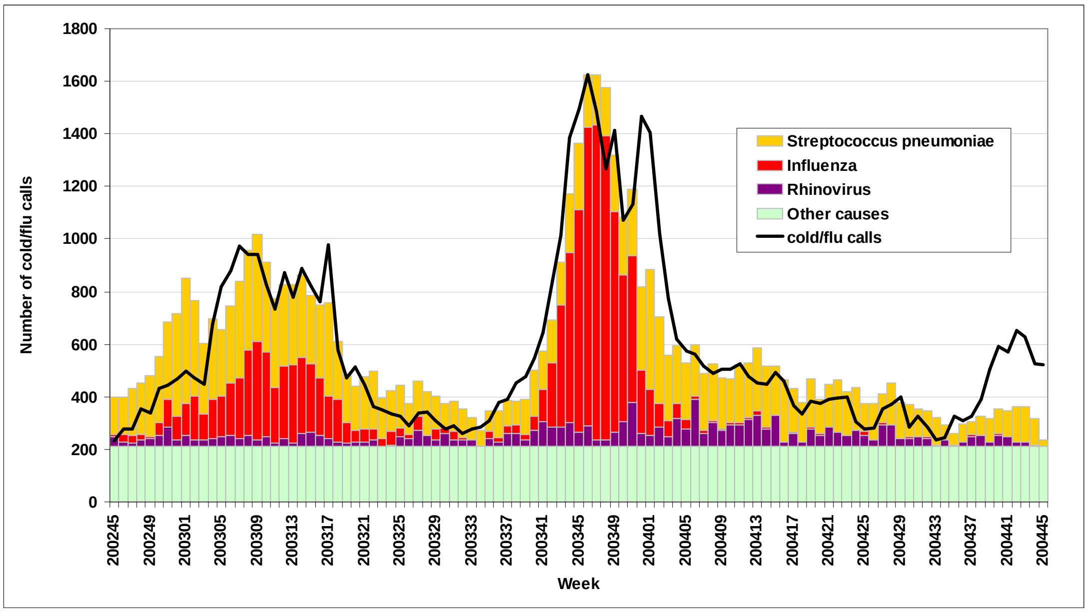


Figure 3a. Comparison of the observed number of NHS Direct cough calls (all ages) with the estimated numbers due to influenza, RSV, *Streptococcus pneumoniae*, rhinovirus, bank holidays and other causes (October 2002 – October 2004), (model C1).

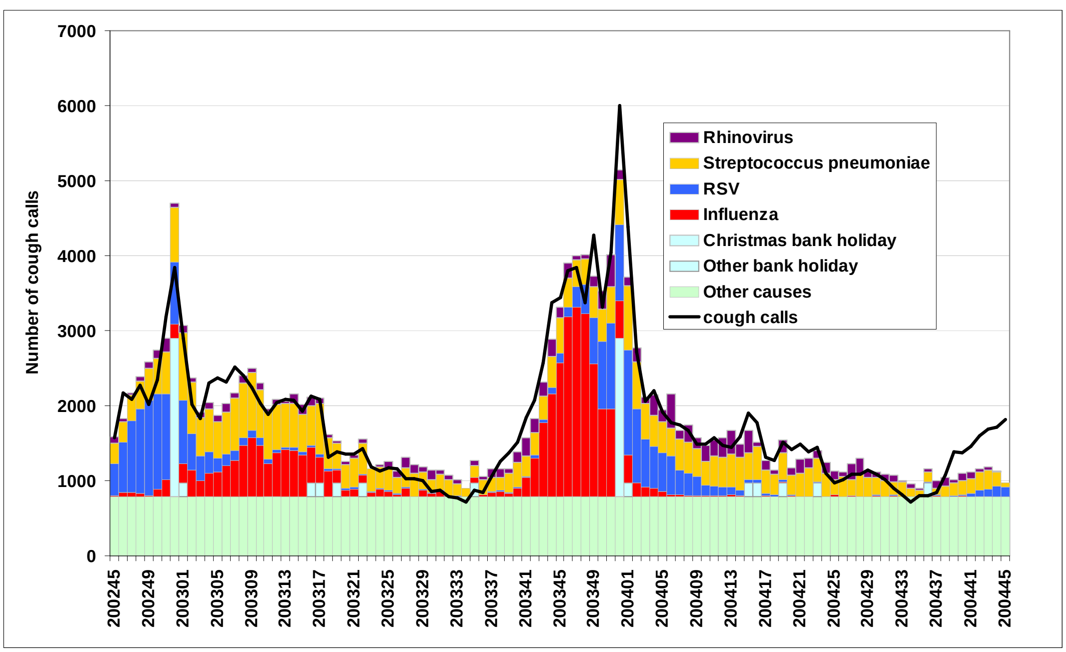


Figure 3b. Comparison of the observed number of NHS Direct cough calls (0-4yrs) with the estimated numbers due to influenza, RSV, *Mycoplasma pneumoniae*, rhinovirus, parainfluenza, bank holidays and other causes (October 2002 – October 2004), (model C1(0-4)).

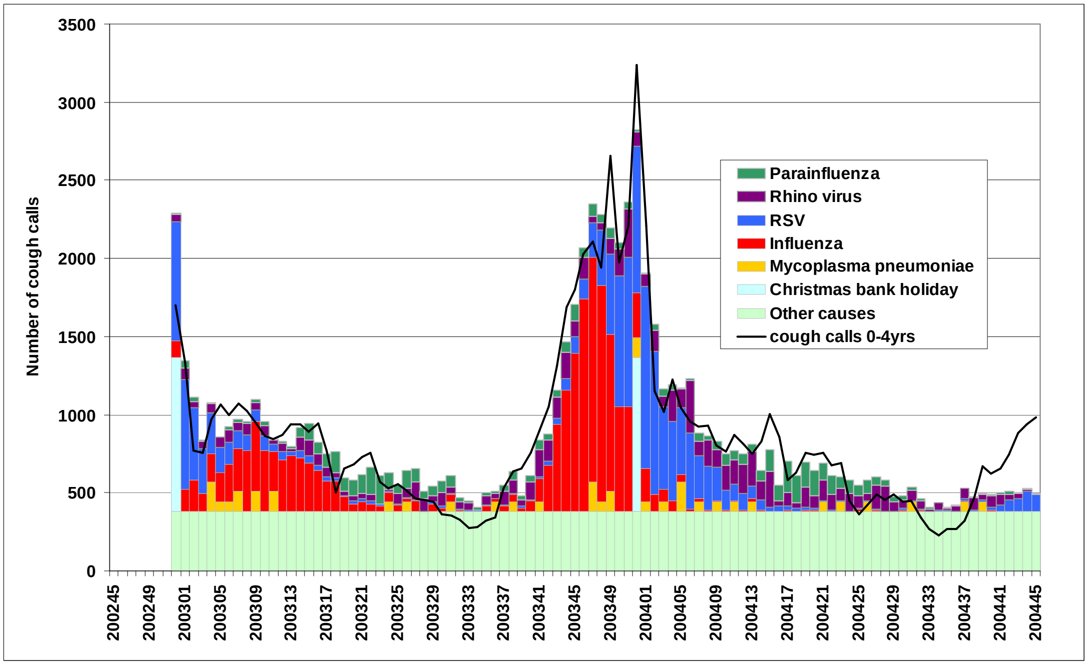


Figure 4. Comparison of the observed number of NHS Direct fever calls (0-4yrs) with the estimated numbers due to influenza, *Streptococcus pneumoniae* and other causes (October 2002 – October 2004), (model F1).

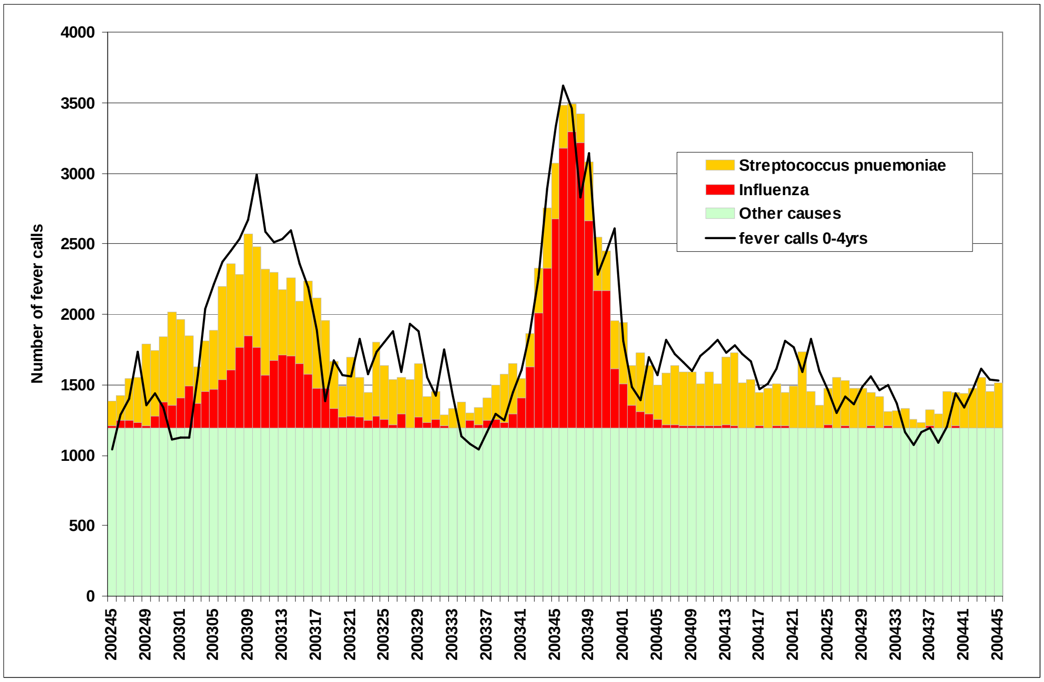


Table 2. Estimated proportion (%) of NHS Direct respiratory calls attributable to specific pathogens for England and Wales (October 2002 – October 2004), and the sensitivity of the model results (R^2).

	Cold/flu calls (all ages)			Cough calls (all ages)			Cough calls (0-4 years)			Fever (all ages)	Fever (0-4 years)	
Pathogen	Model CF1 Final Model	Model CF2	Model CF3	Model C1 Final model	Model C2	Model C3	Model C1 (0-4) Final model	Model C2 (0-4)	Model C3 (0-4)	Model F1 Final model	Model F1 (0-4) Final model	Model F2 (0-4)
Influenza	22	25	26	15	17	-	19	20	23	13	13	14
RSV	-	-	-	12	15	18	16	17	-	-	-	-
Rhinovirus	8	10	-	6	-	-	10	-	-	-	-	-
Parainfluenza		-	-	-	-	-	6	-	-	-	-	-
Streptococcus pneumoniae	33	-	-	20	-	-	-	-	-		19	-
Haemophilus influenzae	-	-	-	-	-	-	-	-	-	19	-	-
Mycoplasma pneumoniae	-	-		-	-	-	3	-	-	-	-	-
Christmas and other bank holidays	-	-	-	3	4	2	2	2	2	-	-	-
Other causes	37	65	74	43	64	80	44	61	75	68	68	86
R^2	0.79	0.72	0.71	0.91	0.87	0.50	0.88	0.86	0.54	0.71	0.76	0.68

Table 3. Estimated annual incidence of NHS Direct respiratory calls due to specific pathogens (per 100,000 with 95% confidence intervals) for England and Wales (October 2002 to October 2004).

	Cold/flu (all ages) Model CF1	Cough (all ages) Model C1	Fever (all ages) Model F1	Total (all ages)	Cold/flu (0-4 years) Model CF1 (0-4)	Cough (0-4 years) Model1 (0-4)	Fever (0-4 years) Model F1 (0-4)	Total (0-4 years)
Influenza	12.8 (11.2-14.5)	27.2 (23.9-30.4)	32.6 (28.3-36.9)	72.6 (63.4-81.8)	16.3 (13.9-18.7)	247.1 (213.3-280.5)	367.5 (321.4-413.9)	630.9 (548.6-713.1)
RSV	-	21.3 (17.2-25.4)	-	21.3 (17.2-25.4)	7.3 (4.6-10.2)	202.9 (161.4-244.4)	-	210.2 (166-254.6)
Streptococcus pneumoniae	19.1 (12.7-25.5)	36.2 (21.6-50.6)	-	55.5 (34.3-76.1)	23.5 (14.1-32.8)	-	542 (360.2-723.8)	565.5 (374.3-756.6)
Haemophilus influenzae	-	-	48.9 (29.6-68.2)	48.9 (29.6-68.2)	-	-	-	-
Mycoplasma pneumoniae	-	-	-	-	-	39.3 (4.8-73.9)	-	39.3 (4.8-73.9)
Rhinovirus	4.4 (0.5-8.1)	11.2 (3.1-19.2)	-	15.6 (3.6-27.3)	-	126.6 (39.8-213.4)	-	126.6 (39.8-213.4)
Parainfluenza	-	-	-	-	-	81.5 (10.7-152.3)	-	81.5 (10.7-152.3)

Estimated incidence

The estimated mean annual incidence of NHS Direct respiratory calls (cold/flu + cough + fever calls) attributable to influenza was 72.6 per 100,000, and to *S. pneumoniae* (cold/flu + cough) 55.5 per 100,000 (Table 3). The estimated incidence of NHS Direct cough calls attributable to RSV was 21.3 per 100,000. NHS Direct respiratory calls about children under 5 years attributable to influenza were estimated to be 630.9 per 100,000, 210.2 per 100,000 for RSV and 565.5 per 100,000 for *S. pneumoniae*.

Discussion

Annual peaks in NHS Direct cough and difficulty breathing calls occurred in late December, whereas peaks in cold/flu and fever calls were less predictable and occurred between November and April. The regression models estimated that at least half of the seasonal variation in NHS Direct respiratory calls was due to the seasonality of influenza and RSV. During short lived winter periods, influenza and RSV accounted for half of the NHS Direct calls about cold/flu, cough and fever. Rhinovirus accounted for a tenth of cold/flu and cough calls, distributed evenly throughout the year, with parainfluenza estimated to account for a quarter of cough calls (0-4 years) during spring. The main bacterial cause of NHS Direct respiratory calls was *S. pneumoniae*, with an estimated year round contribution, peaking during mid-winter. The cough models demonstrated the importance of incorporating a bank holiday factor when modelling NHS Direct call data and provide evidence that closure of GP surgeries during bank holidays stimulates acute demand for NHS Direct: a third of all cough calls during Christmas week were accounted for by this bank holiday effect.

Although a good fit was achieved for the final cold/flu, fever and cough models, methodological problems were encountered. Dropping variables with negative coefficients from the models before re-estimating is an inelegant approach to multiple linear regression. Resulting bias due to this is considered negligible, however, as the removal of these variables did not greatly alter the values of the significant positive coefficients (e.g. RSV, influenza, *S. pneumoniae*) in the final models. The statistical models described here use laboratory data which are subject to case ascertainment bias. For example, common causes of respiratory infections (e.g. coronavirus [13], and human metapneumovirus [14]) that peak during winter, are rarely tested for and are therefore present in low numbers within LabBase. These pathogens do contribute to the intercept (other causes), however, and are therefore implicitly included within our models. Consequently, over estimation of the contribution of significant variables (e.g. RSV, influenza, and *S. pneumoniae*) is considered negligible.

The low sensitivity of the two difficulty breathing models (R^2 = 0.23 and 0.48) may have been improved by adding environmental factors (e.g. pollen, air quality) as explanatory variables. This could also have improved the model estimates of the number of cough calls during October-November 2004, when large model residuals were encountered. Previous work has implicated environmental pollutants in sudden increases in difficulty breathing calls [15] and further work is needed to explore non-microbiological influences on NHS Direct respiratory calls. Not surprisingly, we noted a relatively high estimated proportion of fever calls due to other causes (>60% in models F1 and F1 (0-4)), which are likely to be caused by a variety of exanthema, gastrointestinal illness and other diseases.

NHS Direct currently handles approximately 7 million clinical calls per year compared to 190 million GP consultations [16] and 14 million A&E admissions in England and Wales [17]. The estimates of specific respiratory diseases reported to NHS Direct provide additional baseline data to be placed alongside GP and hospital data when assessing the burden of disease and evaluating the potential impact of vaccination strategies (e.g. for pneumococcal vaccine [18]). The results highlight the importance of influenza, RSV and *S. pneumoniae* infections in respiratory symptoms reported to NHS Direct, supporting previous work showing these diseases also impact substantially on GP consultations [9,19] and hospitalisations [8,14] in the UK. The contribution of *S. pneumoniae* disease to NHS Direct respiratory calls (an estimated 92,000 calls per year; one sixth of total NHS Direct respiratory calls) compares to an estimated 700,000 annual GP consultations (approximately 25% of community acquired pneumonia (CAP) and otitis media) [9], and 29% of patients admitted to hospital with CAP [20] (20% for children under 5 =years [8]). The NHS Direct estimate is likely to be lower because the NHS Direct denominator (total respiratory calls) covers a wider and potentially milder spectrum of disease than the GP denominator (GP CAP and otitis media) and hospital denominator (admissions for CAP).

With regard to influenza, a pilot study has demonstrated that when NHS Direct callers self-sampled for influenza and posted diagnostic samples to a central laboratory over a 3 month winter period, 16% had laboratory confirmed influenza, with positivity rising above 30% during peak weeks [21]. These data support our model estimates (one in three cold/flu calls attributable to influenza during the 'flu season). Perhaps it is not surprising that an adult suffering from what they consider to be an unusually severe respiratory illness (likely to be more than a common cold) will phone NHS Direct. Common respiratory viruses that cause mild to moderate upper respiratory infections had a smaller estimated contribution to NHS Direct calls (e.g. rhinovirus) or were not significant within our models (e.g. adenovirus). Finally, th*e* two winters covering the study period were characterised by low levels of influenza activity [22-23]. The estimated contribution of influenza to NHS Direct respiratory calls, 13-22% during our study, is dependent on epidemiological conditions and would be greater during years with more virulent strains of influenza.

The weekly estimates of the contribution of different pathogens to NHS Direct respiratory calls will provide greater confidence in interpreting sudden rises in calls detected through syndromic surveillance. For example, a sharp rise in cold/flu calls from adults and/or rise in fever calls about children may represent an increase in influenza circulating in the community. A rise in cough calls during November/December may indicate a rise in RSV and also influenza; a rise during January to March suggest only influenza; and a rise during spring may indicate a rise in parainfluenza infections. The use of this intelligence within the NHS Direct syndromic surveillance system will complement other primary care surveillance systems: the Weekly returns Service of the RCGP [10] and the newly established Q-Research system [24]. These pathogen specific estimates may also be of use for NHS winter planners in managing elective hospital admissions and in informing the health service of impending problems. Similarly, analysis of Emergency Department respiratory syndrome data in the US suggests these data could be a useful early indicator of influenza and RSV circulation [25].

This work has provided a detailed weekly picture of the type of respiratory infections - throughout the year – that cause people to phone NHS Direct. This has implications for the advice offered by NHS Direct nurses in handling respiratory calls (currently 8% of total symptomatic calls). The supporting clinical advice and rationales within the decision support system used by NHS Direct nurses (NHS CAS) can be modified to reflect awareness of the prevailing respiratory infections. For example, this would include specific advice on seeking urgent help in the event of RSV being the likely cause of significant acute respiratory impairment in a child (bronchiolitis).

Conclusion

In this report we have described the seasonality of NHS Direct respiratory calls and have used a relatively straightforward modelling technique to estimate the proportions of NHS Direct calls attributable to specific microbiological causes. This suggests that at least half of the variation in NHS Direct respiratory calls is explained by the seasonality of common respiratory viruses, notably influenza and RSV. There is also a significant year round contribution of *S. pneumoniae* infections, peaking during winter. The results of this work will be used to provide detailed estimates of the burden of specific respiratory diseases reported to NHS Direct; advise NHS Direct on the timing and aetiology of respiratory calls; and help interpret syndromic surveillance data used to provide early warning of community rises in respiratory infections.

References

1. Health protection in the 21st Century: Understanding the burden of disease; Preparing for the future. Health Protection Agency 2005.
2. NHS Direct Online. http://www.nhsdirect.nhs.uk/
3. Cooper DL, Smith GE, Hollyoak VA, Joseph CA, Johnson L, Chaloner R. Use of NHS Direct calls for the surveillance of influenza – a second year's experience. *Commun Dis Public Health* 2002; **5**(2): 132-136.
4. Baker M, Smith GE, Cooper DL, et al. Early Warning and NHS Direct: a role in community surveillance? *J Publ Hlth Med* 2004; **25**:362-368.
5. Lazarus R, Kleinman K, Dashevsky I, DeMaria A Jr, Platt R. Using automated medical records for rapid identification of illness syndromes (syndromic surveillance): the example of lower respiratory infection. *BMC Public Health* 2001; **1**: 9. http://www.biomedcentral.com/1471-2458/1/9/
6. Heffernan R, Mostashari F, Das D, Karpati A, Kulldorff M, Weiss D. Syndromic surveillance in public health practice, New York City. *Emerg Infect Dis* 2004; **10**(5): 858-864.
7. Das D, Metzger K, Heffernan R, Balter S, Weiss D, Mostashari F. Monitoring over-the-counter medication sales for early detection of disease outbreaks - New York City. *MMWR* 2005; **54** Suppl:41-46.
8. Muller-Pebody B, Edmunds WJ, Zambon MC, Gay NJ, Crowcroft NS. Contribution of RSV to bronchitis and pneumonia-associated hospitalizations in English children, April 1995-March 1998. *Epidemiol Infect* 2002; **129**: 99-106.

9. Melegaro A, Edmunds WJ, Pebody R, Miller E, George R. The current burden of pneumococcal disease in England and Wales. *J Infect* 2006; **52**(1):37-48.
10. Fleming DM. Weekly Returns Service of the Royal College of General Practitioners. *Commun Dis Public Health* 1999; **2**(2):96-100. Review.
11. Joseph CA, Harrison TG, Ilijic-Car D, Bartlett CL. Legionnaires' disease in residents of England and Wales: 1998. *Commun Dis Public Health* 1999 ;**2**(4):280-4.
12. Ryan MJ, Ramsay M, Brown D, Gay NJ, Farrington CP, Wall PG. Hospital Admissions Attributable to Rotavirus in England and Wales. *J Infect Dis* 1996; **174**(Suppl):S12-18.
13. van Gageldonk-Lafeber AB, Heijnen ML, Bartelds AI, Peters MF, van der Plas SM, Wilbrink B. A case-control study of acute respiratory tract infection in general practice patients in The Netherlands. *Clin Infect Dis* 2005;**41**(4):490-7.
14. Nicholson KG, McNally T, Silverman M, Simons P, Stockton JD, Zambon MC. Rates of hospitalisation for influenza, respiratory syncytial virus and human metapneumovirus among infants and young children. *Vaccine* 2006; **24**(1):102-8. Epub 2005 Feb 25.
15. Cooper DL, Smith GE, Baker M, et al. National Symptom Surveillance Using Calls to a Telephone Health Advice Service — United Kingdom, December 2001– February 2003. *MMWR* 2004*;* **53** (Supp): 179-183.
16. Rowlands S, Moser K. Consultation rates from the general practice research database. Br J Gen Pract 2002; 52:658–60.
17. England Department of Health. Hospital activity statistics. Available at http://www.performance.doh.gov.uk/hospitalactivity/index.htm.
18. Salisbury D, Begg N. Immunisation against infectious diseases. London: HMSO; 1996.
19. Zambon MC, Stockton JD, Clewley JP, Fleming DM. Contribution of influenza and respiratory syncytial virus to community cases of influenza-like illness: an observational study. *Lancet* 2001; **358** (9291):1410-6.
20. Hawker J, Begg N, Blair I, et al. *Communicable Disease Control Handbook.* Oxford, Blackwell Science Limited, 2001: P21.
21. Cooper DL, Smith GE, Loveridge P, et al. Can we use self-testing to augment syndromic surveillance? A pilot study using *influenza. MMWR* 2005*;* 54 (Suppl): 186.
22. Cooke MK, Crofts JP, Joseph CA, et al. Influenza and other respiratory viruses surveillance in the United Kingdom: October 2002 to May 2003. CDR 2004; Suppl: 1-8.
23. Crofts JP, Joseph CA, Zambon M, et al.. Influenza and other respiratory viruses surveillance in the United Kingdom: October 2003 to May 2004. CDR 2003; Suppl: 1-9.
24. Harcourt S, Smith GE, Hippisley-Cox J, et al. Report of the first year of a pilot, national, primary care surveillance project. *Proceedings of the Health Protection Agency Annual Conference* 2005; P 119.
25. Bourgeois FT, Olson KL, Brownstein JS, McAdam AJ, Mandl KD. Validation of syndromic surveillance for respiratory infections. *Ann Emerg Med* 2006 Mar;**47**(3):265.e1.

8 Tracking the spatial diffusion of influenza and norovirus using telehealth data: a spatio-temporal analysis of syndromic data

Submitted as: Cooper DL, Smith GE, Regan M, Large S, Groenewegen PP. Tracking the spatial diffusion of influenza and norovirus using telehealth data: a spatio-temporal analysis of syndromic data. *BMC Medicine 2008; 6:16.*

Abstract

Background: Telehealth systems have a large potential for informing public health authorities in an early stage of outbreaks of communicable disease. Influenza and norovirus are common viruses that cause significant respiratory and gastrointestinal disease world wide. Data about these viruses are not routinely mapped for surveillance purposes in the UK; so the spatial diffusion of national outbreaks and epidemics is not known as such incidents occur. We aim to describe the geographical origin and diffusion of rises in fever and vomiting calls to a national telehealth system, and consider the usefulness of these findings for influenza and norovirus surveillance.

Methods: Data about fever calls (5-14yrs) and vomiting calls (≥5yrs), proxies for influenza and norovirus respectively, were extracted from the NHS Direct national telehealth database for the period June 2005 to May 2006. The SaTScan space-time permutation model was used to retrospectively detect statistically significant clusters of calls on a week by week basis. These syndromic results were validated against existing laboratory and clinical surveillance data.

Results: We identified two distinct periods of elevated fever calls. The first originated in the North West of England during November 2005 and spread in a south-east direction, the second began in Central England during January 2006 and moved southwards. The timing, geographical location, and age structure of these rises in fever calls were similar to a national influenza B outbreak that occurred during winter 2005/2006. We also identified significantly elevated levels of vomiting calls in South-East England during winter 2005/2006.

Conclusion: Spatio-temporal analyses of telehealth data, specifically fever calls, provided a timely and unique description of the evolution of a national influenza outbreak. In a similar way the tool may be useful for tracking norovirus, though lack of consistent comparison data make this more difficult to assess. In interpreting these results, care must be taken to consider other infectious and non-infectious causes of fever and vomiting. The scan statistic should be considered for spatial analyses of telehealth data elsewhere and will be used to initiate prospective geographical surveillance of influenza in England.

Background

Disease surveillance systems are used for identifying important events and trends within the distribution of known diseases. This information should inform and drive public health action. Traditionally laboratory reports and clinical diagnoses are used for disease surveillance, providing epidemiological details about the time, person and place associated with disease. In recent years there has been a growth in syndromic surveillance systems that collect and analyse pre-diagnostic data to provide early warning of infectious disease outbreaks or potential bio-terrorist attacks [1]. These systems have been successful in detecting rises in syndromes associated with common viral diseases [2, 3]. Influenza and norovirus are two such viruses that cause significant respiratory and gastrointestinal disease in the UK [4,5] and world wide. Both are responsible for outbreaks within health-care [6] and educational settings [7], and are associated with significant economic costs [8, 9].

The influenza virus causes an acute infection of the respiratory tract characterised by cough, fever and myalgia, and in more serious cases can lead to pneumonia. Its epidemiology has been studied in great detail, partly due to the ability of the virus to mutate causing national epidemics and worldwide pandemics (e.g during 1918, 1957 and 1968). On a population level the disease spreads via a combination of contagious diffusion (wave like from a central or multiple foci) and hierarchical diffusion (movement from large to smaller towns) [10]. During the 1918-1919 pandemic in the UK influenza moved in a southerly, then northerly, then southerly direction during three successive epidemic waves [11]. Although there has been a secular decline in the incidence of influenza-like-illness (ILI) in the UK in recent decades [4] elevated activity still occurs during winter, often reaching peak levels first in Northern England [7, 12]. For example, the winter of 2005/2006 was marked by a national outbreak of influenza B and nearly 700 reported school outbreaks during January and February. A well established surveillance programme [13] published weekly reports throughout the winter, summarising clinical incidence and laboratory data and tabulating regional levels.

Norovirus is associated with projectile vomiting and diarrhoea. There are an estimated 700,000 cases in England each year [5], although it is estimated that only one out of every 1600 cases reach national laboratory surveillance [14], prompting calls for methods of supplementing current surveillance systems. Like influenza, new strains of norovirus may emerge [15] causing significant illness and a rise in reported outbreaks. The geographical variation in community infection rates of norovirus is poorly understood in the UK as the 'gold standard' for surveillance is considered to be outbreak reports. These reports originate mainly from institutional rather than community settings [6] and are neither timely nor reported in a consistent manner across the country.

Neither influenza nor norovirus data are routinely mapped for surveillance purposes in the UK. This means that the local incidence, and the spatial diffusion of national epidemics cannot be described prospectively for public health teams, health service planners and the public. The Health Protection Agency currently collects and analyses data from a national telephone triage service (NHS Direct) for syndromic surveillance purposes [16]. NHS Direct covers the whole of England and Wales and is widely used by the population. NHS Direct nurses record demographic and syndromic call details over the telephone so the call data offer potential for disease surveillance. Therefore, analysing the usefulness of NHS Direct data has a wider relevance for public health and preparedness.

In this article we have used syndromic data about NHS Direct fever and vomiting calls, of which there are many infectious and non-infectious causes. On a national and regional level, NHS Direct fever calls for school age children rise at the same time as laboratory reports of influenza B [17]. Vomiting calls rise during periods of increased norovirus activity nationally [18]. Developments in NHS Direct data management systems means that syndromic call data can now be extracted and analysed to a local level. In this study we aimed to characterise the geographical origin and diffusion of rises in NHS Direct fever and vomiting calls (i.e. potential epidemics). Freely available spatio-temporal analyses software (SaTScan) was used to detect clusters of elevated NHS Direct calls on a week by week basis. We compared these syndromic results to existing laboratory and clinical surveillance data in order to determine whether NHS Direct data provide a sensitive and timely means of describing local and national level increases in common viruses.

Methods

Data collection

Data were extracted from the NHS Direct national database about the syndrome, time of call, age, and postcode district (geographical identifier) of each NHS Direct call in England for the period June 2005 to May 2006. At the time of analysis a single year's postcoded data were available. A geographical information system (MapInfo GIS) was used to assign a Primary Care Trust (PCT) code to each call. The PCT is the local unit of administration for primary and public health services in England. There are 151 PCTs with an average population of 353,000 people. Two subsets of NHS Direct data were extracted; fever calls for the 5-14 year age group, used as a proxy for influenza B infection which predominantly affects school aged children; and vomiting calls for the ≥5 year age group used as a proxy for norovirus infection. Vomiting calls about those <5yrs were excluded from the analysis to remove the potentially confounding effect of rotavirus (another common viral cause of vomiting which almost exclusively affects children under 5). It is not currently possible to extract multiple symptoms linked to individual callers due to the way in which call details are stored within the NHS Direct call database.

Statistical methods

Data were analysed using the SaTScan scan statistic to analyse spatial, temporal or spatial-temporal data [19]. The scan statistic uses circular windows (potential cluster areas) starting as a single point (PCT location) and increasing in radius to form an additional circle each time a new point is reached. The maximum spatial extent and time period of the circles is defined by the user; in this case it was set to extend to 50% of the total population of England. For each circle the number of observed and expected cases is calculated as well as a likelihood ratio and relative risk to determine whether the call rate inside the circle is greater than the area outside (i.e. a cluster). The statistical significance of clusters with an excess of cases (observed > expected) is calculated by using Monte Carlo hypothesis testing to compare the real data against randomly generated datasets. Clusters with a significantly elevated relative risk ($p \leq 0.05$) were stratified into 5 levels of risk and mapped using a GIS. If a PCT was contained within two overlapping clusters the relative risk with the lowest p value was mapped, a method borrowed from Boscoe to visualize SaTScan output [20]. The application of the scan statistic to communicable disease surveillance has been described extensively elsewhere [21-24].

Analyses

We used a three stage process to address our aims. For stage 1 we mapped annual total NHS Direct call rates by PCT and used the SaTSCan scan statistic to identify areas of high call rates all year round ('global clustering'). Global clustering may bias the results of the weekly analyses (stage 2) by causing frequent but misleading clusters of syndromic calls in areas of high total call rates. Accounting for the results of stage 1, stage 2 was a spatio-temporal analysis of syndromic calls to describe the location and diffusion of significant rises in fever and vomiting calls. Stage 3 used existing influenza and norovirus surveillance sources (GP consultations, laboratory reports, outbreaks) to describe the accuracy and usefulness of NHS Direct syndromic data. A similar staged approach has been used previously to identify clusters of E Coli 0157 cases in Canada [24].

Table 1. Numbers, rates per 1000/year and percentages of syndromic calls (June 2005 – May 2006).

	Number of calls	*Call rate per 1000/year (range within PCTs)*	*Syndromic calls, as a percentage of total calls for same age-group*
Vomiting calls (≥5yrs)	88,452	1.8 (0.6-3.2)	2.5%
Fever calls (5-14yrs)	23,431	3.7 (1.1-8.5)	6.4%

Stage 1: Mapping total NHS Direct call rates

Total annual NHS Direct call rates (all ages, 5-14yrs, and ≥5yrs) were mapped by PCT. We then conducted a single retrospective spatial analysis of one years call data using the SaTScan Poisson probability model. This tested for areas with significantly high or low annual call rates over the whole study period. Three data files were required for analysis; the number of calls per week in each PCT (numerator), the population of each PCT (denominator), and the grid coordinates of the centroid of each PCT (location). The PCT was preferred as the unit of analysis (rather than the smaller postcode district) to test provision of PCT based data for seasonal influenza surveillance and pandemic influenza preparedness, and to maintain statistical stability by avoiding generating clusters with very small numbers of calls which in practice are unlikely to be investigated further. This choice of geographical unit is discussed later in the article.

Stage 2: Spatio-temporal analyses of syndromic calls

The results of stage 1 (described in the results section) indicated that there was significant global clustering and spatial heterogeneity in local call rates. To correct for this underlying spatial variation, the SaTScan space-time permutation model was used to conduct further analyses. This approach requires only case data (in this case either fever or vomiting calls) and adjusts for purely spatial or purely temporal clusters. For this method the number of observed calls in a cluster is compared to the expected number, assuming that the spatial and temporal locations of calls are independent of each other (no space-time interaction). At each time interval specified, a cluster is identified if a specific area has a higher proportion of excess cases than surrounding areas. Therefore areas with historically high usage of NHS Direct (identified in stage 1) will not bias the results. Also, if during a specific week (e.g. the Christmas holiday when doctor's surgeries are closed) all areas experience a doubling of vomiting calls, no clusters will be identified. Such a national rise would be detected by the existing NHS Direct syndromic surveillance system [16]. Changes in PCT population over the study period could hypothetically cause clusters in areas with increased population and reduce the probability of clusters in areas with decreased population. Our study period covered a single year so population change is considered negligible.

We used the space-time permutation model (on historical data) to mimic prospective surveillance and identify 'live' clusters (clusters ending with the most recent date) lasting 1 week and with a p value of ≤ 0.05. We retained the maximum cluster population size of 50%. SaTScan uses cases (calls) as a proxy for population in the space-time permutation model as only case data are included. The first 4 months data (June 1st 2005 – 2nd October 2005) were used as the control period on which scan statistic calculates an expected count for each location. Repeated analyses were performed for each additional week's data for a test period from October 3rd 2005 to May 21st 2006, for fever calls (5-14yrs) and vomiting calls ($\geq$5yrs) separately. Although daily NHS Direct data were available we conducted weekly analyses to simplify presentation of data and align temporal comparisons with clinical and laboratory data which are available weekly. Details of the location, observed and expected calls, and level of significance of each significant space-time cluster were recorded. The ratio of observed to expected calls (O/E) for each cluster was mapped for each weekly analysis using the same method as in stage 1.

Stage 3: Comparison against existing surveillance data

The dates and spatial distribution of fever clusters (identified in stage 2) were compared against routinely available influenza surveillance data; weekly influenza-like-illness (ILI) rates for the 5-14yr age group for North, Central and Southern England (obtained from the Royal College of General Practitioners Weekly Returns Service) [25], and the weekly number (all ages combined, infleunza A and B separately) of positive influenza samples from community sources from the Health Protection Agency, Centre for Infections virological surveillance scheme [13]. The dates and spatial distribution of vomiting clusters were compared against weekly numbers of norovirus outbreaks recorded by HPA Centre of Infections Modular Open Laboratory System (MOLIS), presented by the region of origin of laboratory samples. The MOLIS database was recommended as a suitable source of outbreak data as each laboratory sample has an outbreak code and report date making it possible to derive the number of outbreaks per week and region.

Results

During the study period there were 4,728,939 NHS Direct calls from the public seeking advice for all symptoms of which 4,406,607 (93%) were linked via a postcode to a PCT. The total call rate was 82.1 calls per 1,000 people, ranging from 33.9 to 212 per 100,000 between PCTs. Annual syndromic calls for fever (5-14yrs) and vomiting (≥5yrs) are summarised in Table 1.

Stage 1

The highest all symptom annual call rates (in excess of 100 calls per 1000/year) were found particularly in the North West and Yorkshire and Humber regions, with other areas of high incidence areas scattered throughout England (Figure 1). Spatial analyses of these data highlighted two areas of significantly high call rates (relative risk (RR): 1.3-1.5, p=0.001) in Northern England and Nottinghamshire, and a widespread area of significantly low call rates across Central and Eastern England

(RR: 0.8-0.9, p=0.001) (Figure 2). Total call rates for the 5-14yr and ≥5yr age groups exhibited a similar geographical variation by PCT (not presented).

The significant geographical variation in total call rates meant that if we had adopted a similar SaTScan model for the iterative weekly analyses of syndromic data (stage 2) the results would have been largely determined by this underlying variation in call rates. The space-time permutation model was used, therefore, whereby a baseline period is used to establish the expected weekly number of syndromic calls in each PCT.

Stage 2

There were 27 fever clusters during the study period (p≤0.01), all but one within a 15 week period between November 2005 and March 2006 (Table 2). The clusters ranged from 2 to 72 PCTs in size (mean 34), and from 18 to 937 fever calls (mean 310). The shaded areas in the maps (Figure 3) denote areas with statistically high observed numbers of fever calls ('clusters') compared with the expected number. The first significant cluster was detected in North West England during week 2005/47. This area of raised fever incidence increased in size to cover large areas of Central and Northern England by week 2005/52. A second rise in fever calls began with a cluster in Central England during week 2006/02 and spread south, reaching its maximum extent during week 2006/07. Isolated clusters lasting a single week were detected in Hertfordshire (week 2005/50) and Northumberland (week 2006/14).

There were 22 vomiting clusters (p≤0.01), all but three within a 14 week period between December 2005 and March 2006 (Table 3). The clusters ranged from 4 to 74 PCTs in size (mean 32), and from 85 to 1345 vomiting calls (mean 555). Sporadic 1-week clusters of vomiting calls were observed in the North West and Southern England between Weeks 2005/40 and 2005/50. A widespread significant rise in vomiting calls occurred in South Eastern England in week 2005/51 reaching its maximum extent in week 2006/05.

Stage 3

The pre-Christmas significant cluster of NHS Direct fever calls (5-14yrs) in North West England (weeks 47 and 48 of 2005) occurred prior to a pre-Christmas rise in GP diagnosed influenza-like illness (ILI) (5-14yrs) in Northern England (peaking week

50/2005) and reports of early influenza B outbreaks in primary schools in the North West reported in week 49/05 [7]. The second significant rise in NHS Direct fever calls originated in Central England during week 02/06, two weeks prior to a rise in ILI above baseline levels in Central England (30 per 100,000 [4]). The subsequent southerly movement and expansion of this rise in fever calls during weeks 03-07/2006 occurred concurrently with a rise in ILI peaking firstly in Central England (week 06/06) and then Southern England (week 07/06), and a rise in laboratory reports of influenza B peaking in weeks 05-06/06. A rise in laboratory reports is defined as five or more laboratory report during two consecutive weeks [4]).

The majority of the norovirus outbreaks from the MOLIS database were from the North West and London and South East Regions (84%). There was a period of elevated numbers of outbreaks in the London and South East Regions between November December 2005 and March 2006 (week 02/06 peak = 26 outbreaks) and a later peak in North West outbreaks during April (week 15/06 peak = 34). The period of significantly high vomiting calls in South East England was contained within this period (December 2005 to mid-February 2006) with localised vomiting clusters in South East England lasting until March 2006.

Discussion

Spatio-temporal analyses of telehealth calls may serve as a useful adjunct to influenza surveillance. Our analyses of NHS Direct fever calls provided a clear and detailed geographical description of seasonal influenza in England in 2005/06. We identified two distinct periods of significantly high NHS Direct fever calls about school aged children. The first rise originated in the North West of England during late November 2005 and spread in a predominantly eastwards and then southerly direction over the following month. The second rise began in Central England during mid-January 2006 and moved southwards to eventually cover all of southern England. The timing and geographical location of these rises in fever calls appear similar to a national influenza B outbreak occurring during winter 2005/2006 affecting the same age group, although they were visible slightly earlier than in other data sources. We also identified significantly elevated numbers of vomiting calls in Central and South-East England during December 2005 to February 2006, within a longer period of elevated norovirus outbreak reports in the South East. There was no apparent spatial correlation between vomiting calls and an April 2006 rise in norovirus outbreaks in the North East of England. Assessment of the usefulness of vomiting call data for monitoring the spread of norovirus was limited by inconsistent regional level outbreak sample data within our norovirus data source.

Table 2. Description of all significant fever clusters (p≤0.05) detected by the scan statistic during our test period (October 23rd 2005 to May 21st 2006)

Year/ Week	Location (centroid PCT)	Number of PCTs involved (radius in km)	Observed calls (O)	Expected calls (E) [based on control period]	O/E	p-value
2005/47	Ashton, Leigh and Wigan	8 (20)	34	15	2.23	0.005
2005/48	Warrington	8 (22)	42	20	2.15	0.002
2005/49	North Lancashire	11 (61)	71	31	2.27	0.001
2005/49	Salford	19 (39)	100	57	1.75	0.001
2005/50	East Lancashire	29 (72)	214	125	1.71	0.001
2005/50	East and North Hertfordshire	2 (22)	18	5	3.33	0.001
2005/51	North Lancashire	50 (147)	394	245	1.6	0.001
2005/51	South Staffordshire	59 (125)	404	309	1.31	0.01
2005/52	Wirral	22 (71)	156	100	1.56	0.001
2005/52	Hull	72 (185)	347	278	1.25	0.008
2006/02	Derbyshire County	12 (50)	114	66	1.74	0.001
2006/03	Nottingham City	21 (75)	266	170	1.56	0.001
2006/03	Telford and Wrekin	20 (74)	191	135	1.42	0.001
2006/04	Leicester City	31 (92)	405	276	1.47	0.001
2006/04	Shropshire County	15 (64)	175	110	1.59	0.001
2006/04	Bath and North East Somerset	31 (140)	325	251	1.29	0.002
2006/05	Gloucestershire	76 (154)	937	722	1.30	0.001
2006/05	Mid Essex	72 (185)	879	720	1.22	0.001
2006/06	Southampton city	69 (164)	889	661	1.34	0.001
2006/06	North East Essex	66 (187)	855	660	1.30	0.001
2006/07	Torbay	32 (250)	337	251	1.34	0.001
2006/07	East Sussex Downs and Weald	68 (194)	549	453	1.21	0.003
2006/08	West Sussex	18 (64)	117	72	1.61	0.001
2006/08	Oxfordshire	38 (87)	201	149	1.35	0.01
2006/09	West Hertfordshire	21 (34)	124	74	1.68	0.001
2006/09	Berkshire West	30 (76)	212	148	1.43	0.001
2006/14	Northumberland	6 (62)	33	13	2.46	0.001

We will to discuss these results from two perspectives. The first and general perspective relates to the potential of a telephone triage system, such as NHS Direct, for surveillance purposes; the second perspective for discussion concerns the specific findings and limitations of the current spatial analysis.

The potential of telephone triage systems for surveillance

Telephone triage health systems may be useful for surveillance purposes for a number of reasons. These services are generally available to the entire population of the area covered and, in the case of NHS Direct, are centrally managed. This means that clinical algorithms are chosen in a consistent manner between call centres (although experienced NHS Direct nurses may be more likely to provide self-care advice than more junior colleagues [26]). Secondly, the data are available on a daily basis as opposed to GP and laboratory based surveillance systems which generally report weekly. Thirdly, telehealth is many people’s first or only point of contact with the health service, providing an early opportunity to identify an increase in illness potentially ‘under the radar’ of other systems.

Despite the above, call data will only provide meaningful intelligence if callers reported illness is representative of community morbidity. In England and Wales, total NHS Direct call rates (82 per 1,000 in England) are approximately a thirtieth of annual GP consultation rates for all illness (2,700 consultations per 1,000) [A Elliot personal communications]. One quarter of the population have ever used NHS Direct [27]. Our work identified areas of significantly high NHS Direct call rates in Northern England and substantial variation between PCTs (ranging from 33 to 212 calls per 1000 people per year). This disparity is influenced by a combination of factors including local arrangements for provision of GP out-of-hours services (generating on average 160 calls per 1,000 per annum [28]), social deprivation [29], and the length of time the local NHS Direct site has been in operation [27].

Table 3. Description of all significant vomiting clusters ($p \leq 0.05$) detected by the scan statistic during our test period (October 23rd 2005 to May 21st 2006).

Year/ Week	Location (centroid PCT)	Number of PCTs involved (radius in km)	Observed calls (O)	Expected calls (E) [based on control period]	O/E	p-value
2005/40	Portsmouth City Teaching	11 (81)	229	171	1.34	0.006
2005/46	East and north Hertfordshire	9 (33)	121	80	1.52	0.002
2005/49	Suffolk	31 (113)	367	292	1.26	0.004
2005/50	Luton	4 (22)	85	47	1.79	0.001
2005/51	Bedfordshire	5 (37)	147	86	1.7	0.001
2005/51	Hounslow	41 (59)	811	687	1.19	0.001
2005/52	Eastern and Coastal Kent	59 (167)	1354	1204	1.12	0.007
2006/01	Hammersmith and Fulham	43 (64)	744	612	1.22	0.001
2006/02	Suffolk	36 (115)	525	412	1.27	0.001
2006/03	Bexley	56 (113)	896	737	1.21	0.001
2006/04	Great Yarmouth and Waveney	29 (165)	429	331	1.29	0.001
2006/04	Milton Keynes	55 (93)	799	688	1.16	0.009
2006/05	South Gloucestershire	65 (161)	1023	860	1.19	0.001
2006/05	Brighton and Hove City	46 (97)	789	662	1.19	0.001
2006/05	Cambridgeshire	74 (144)	1177	1046	1.12	0.009
2006/06	Southampton city	44 (116)	822	689	1.19	0.002
2006/06	Bedfordshire	26 (67)	480	394	1.22	0.007
2006/07	East and North Hertfordshire	18 (41)	293	214	1.37	0.001
2006/09	Bedfordshire	11 (52)	249	183	1.36	0.002
2006/10	East and north Hertfordshire	12 (37	241	164	1.47	0.001
2006/10	Newcastle	8 (46)	156	108	1.45	0.002
2006/12	Great Yarmouth and Waveney	31 (167)	469	370	1.27	0.001

NHS Direct receives a disproportionately high number of calls about children under 4yrs (four times higher than the total call rate) and women of child bearing age (twice the male call rate). Both these groups traditionally have a high health service utilisation in the UK [30]. This reporting tendency could have advantages for timely sentinel surveillance of children, to provide early warning of more widespread population impacts. Older people, who have high GP and hospitalisation rates, use NHS Direct the least of any age group. This age and gender bias is consistent with large public telephone triage systems in New Zealand [31], Australia [32] and Canada [33]. NHS Direct data are, therefore, most suited to surveillance of common illnesses in the UK for those aged below 65 years. Measures must be taken to account for spatial variation in usage, such as establishing local baselines from which to detect real variation in call incidence (as employed in this study using a scan statistic).

Prospectively NHS Direct teletriage data have been used to identify sharp rises in syndromes at a regional and national level [16-17]. The retrospective analysis, reported here, has demonstrated that it is also possible to identify a sub-regional rise in calls suggestive of influenza infection. Other potential benefits include the use of the data to inform the NHS Direct responsive messaging service (RM), an operational tool designed to help manage service demand and to inform patients of relevant topical information (e.g. 'flu and colds', 'diarrhoea and vomiting'). Trends in calls may act either as a trigger to initiate responsive messages or provide statistical support for the ongoing provision of health advice. Additionally, as well as monitoring calls suggestive of infections, these data are also used for the acute response to major incidents. For example, during July 2007, daily surveillance of calls identified a statistically significant 40% rise in calls in the Gloucester flood region (caused in part by a rise in people seeking health information and advice). This information was used to brief a national incident team and the Government on the health effects of the floods.

A qualitative evaluation has explored the usefulness of these telephone triage data for surveillance amongst a sample (n=91) of users of the system (those that receive surveillance bulletins and 'alerts') [34]. The most commonly cited benefits of the surveillance were; providing real-time information for incident teams, local Primary Care Trusts, media messages, and question and answer briefings for GPs and the public, supporting national plans (e.g. surveillance of heat stroke calls during the heatwaves of 2003 and 2006), and preparing laboratory staff for an expected increase in specimens.

In recent years, the number of telephone triage systems has increased internationally in reaction to demands for rapid assessment and reassurance of people with health problems, and in an attempt to reduce the use of face to face health services, especially out-of-hours. Therefore, our results may be useful to other countries where there is a perceived need for real-time monitoring of common syndromes (both infectious and non-infectious). The ability of the teletriage reporting software to map call data in real-time (not the case with NHS Direct) would also prove useful for analysing local demand for telephone triage, and for evaluating the impact of promotional activities.

Strengths and weaknesses of the current spatial analysis.

There are several weaknesses to this work. NHS Direct vomiting calls are likely to be caused by a range of viral and bacterial disease pathogens. This may partially hide, within the data, the true seasonality of norovirus reported to NHS Direct. It is possible that only new norovirus variants will impact substantially on the national burden of vomiting calls for this age group (≥5 years) displaying a pattern of national diffusion rather than regional raised incidence. For example, the novel norovirus genogroup II4 variant caused an increase and atypical summer peak in outbreaks across the UK and continental Europe [15] during 2002. Also, norovirus infected individuals phoning NHS Direct may be reflective of norovirus outbreaks in community settings (e.g. schools, food outlets) which exhibit little seasonal variation [6] and therefore do not exhibit clear seasonality within our telehealth data. Conversely, outbreaks in semi-closed institutional settings (e.g. hospitals, residential homes) show a clear winter peak but these cases are unlikely to telephone NHS Direct because they are already receiving care. This work has not studied the spatio-temporal variation in vomiting calls about young children (in which there is the highest incidence of norovirus infection [5]). Calls about this group were excluded to remove the potentially confounding affect of rotavirus which peaks during March in the UK. A study of the epidemic behaviour of rotavirus in the US demonstrated an annual South-West to North-East movement across the country [35]. Analysis of NHS Direct vomiting and diarrhoea calls about young children is required to explore further the spread of viral gastroenteritis in the UK, and respond to calls to supplement existing surveillance systems for infectious intestinal disease [14]. Unfortunately the norovirus outbreak data in our study does not represent outbreaks nationally and data from alternative laboratory or outbreak reporting systems will be sought for future work.

Common respiratory virus's, such as respiratory syncytial virus (RSV), may have contributed to the rise in fever calls (although RSV predominantly effects children less than 5 years). Previous work estimating the contribution of a range of respiratory pathogens to NHS Direct calls found that RSV was responsible for approximately 15% of NHS Direct 'cough' calls [36]. Although a significant relationship was found between the seasonal variation in fever calls and influenza, no significant relationship was observed with RSV, parainfluenza or rhinovirus.

These analyses used only 151 relatively large spatial units (PCTs) to identify the approximate areas of rises in fever calls rather than local disease transmission or hierarchical diffusion of disease from large to smaller urban centres. The lack of spatial precision means these results are subject to the modifiable areal unit problem (MAUP) [37] whereby arbitrary units are used for spatial reporting. Our analysis suffers from what Armhein called the 'scale effect' of the MAUP [38]; the grouping of small areas into larger ones (postcode districts into PCTs). This had the effect of reducing the number of spatial units, increasing the average number of cases within each unit, but reducing the variation in incidence between units. We may have detected significant clustering at a sub-PCT level by using the smaller postcode district, although there were only 0.7 vomiting and 0.3 fever calls per week per postcode district during our study period. It is yet to be demonstrated that syndromic surveillance systems can *prospectively* and *consistently* provide early warning of localised outbreaks of disease. Their utility continues to be seen on a city wide or regional basis at best. This study has used syndromic (proxy) measures of influenza and norovirus to identify a syndromic pattern consistent with contagious influenza diffusion at a PCT level, rather than onward transmission from an index case or single outbreak. PCTs are a relevant spatial unit for syndromic surveillance as primary care is likely to be among the first services to experience the effects of epidemic outbreaks.

The scan statistic employed here has previously been used for analyses of specific infections (e.g shigella [22] and E Coli [24]), ongoing public health surveillance (e.g New York City [2], and monitoring disease during major sporting events [e.g Kentucky Derby Festival [39]). Our work is novel in the use of the scan statistic for surveillance of telehealth calls, a data source that is infrequently used for syndromic surveillance purposes [40]. Although many NHS Direct calls reflect self-limiting disease (e.g. influenza B, norovirus) this illness is still disruptive on a societal level and may go unreported by other health care surveillance systems. This work, for the first time, provides evidence that it is possible to describe the diffusion of national influenza outbreaks using telehealth data. When using the scan statistic the underlying spatial variation in incidence is unimportant until it reaches a significant threshold level, clusters are detected, and an alert may be issued to public health teams. Other statistical approaches to deriving these critical values have been described [41-42]. An alternative approach, however, is kriging where the entire range of disease incidence across a geographical area is modelled. Kriging uses interpolation techniques to estimate unknown point values of disease incidence from surrounding known point values, from which disease contours are generated [43]. The method has been successfully employed to describe the spread of the peak in influenza across Japan [44] and Europe [45]. For prospective surveillance, however, peak weeks are not known in advance and any global spatial clustering caused by differences in case definitions or reporting behaviour between regions would produce misleading results. The space-time scan statistic accounts for these spatial differences by using a control period from which to establish local baselines for each spatial unit. It may therefore be considered for future national surveillance initiatives e.g. the proposed hospital based surveillance and 2012 Olympic planning in the UK.

Figure 1. Annual total call rate per 1,000 population per year (June 2005 – May 2006) mapped by Primary Care Trust.

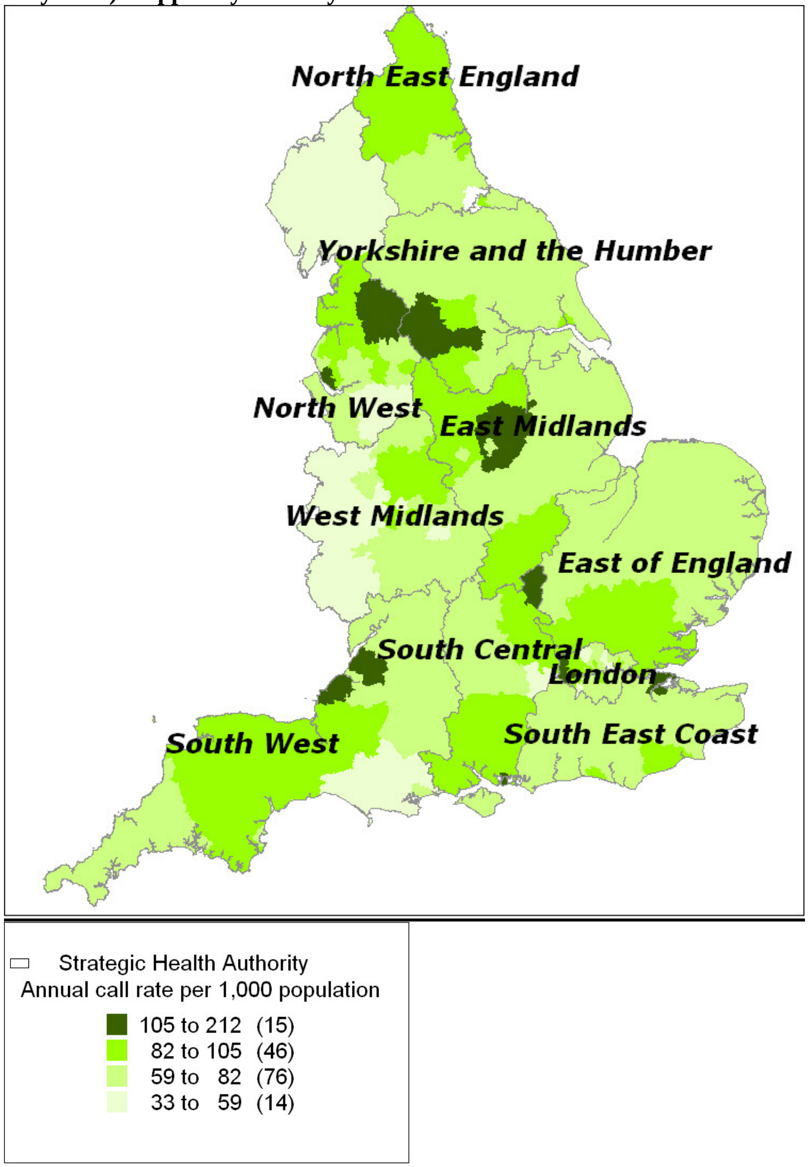


Figure 2. Areas of significantly high or low annual total call rates (June 2005 – May 2006) displayed as relative risks and mapped by Primary Care Trust.

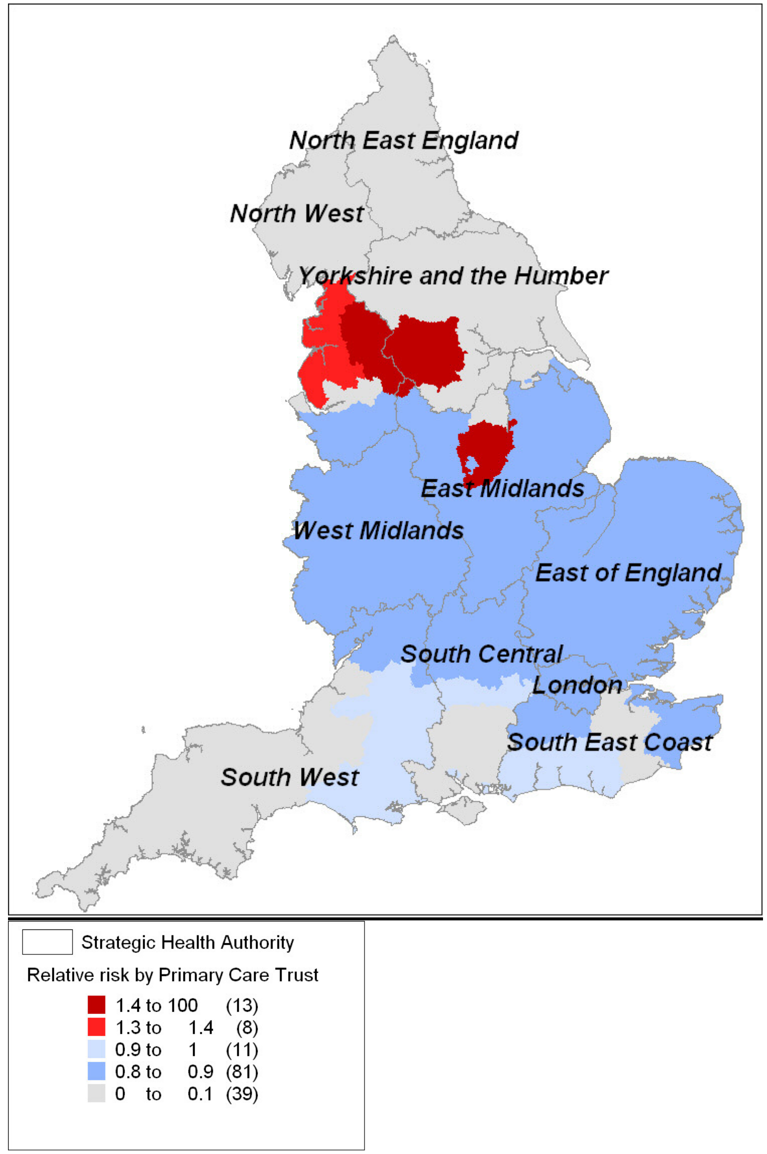


Figure 3. Areas with significantly high numbers of fever calls (clusters) displayed as observed/expected ratios by week and location of the first reported influenza B outbreaks. There were no significant clusters prior to 2005/47 and after 2006/14 during our test period.

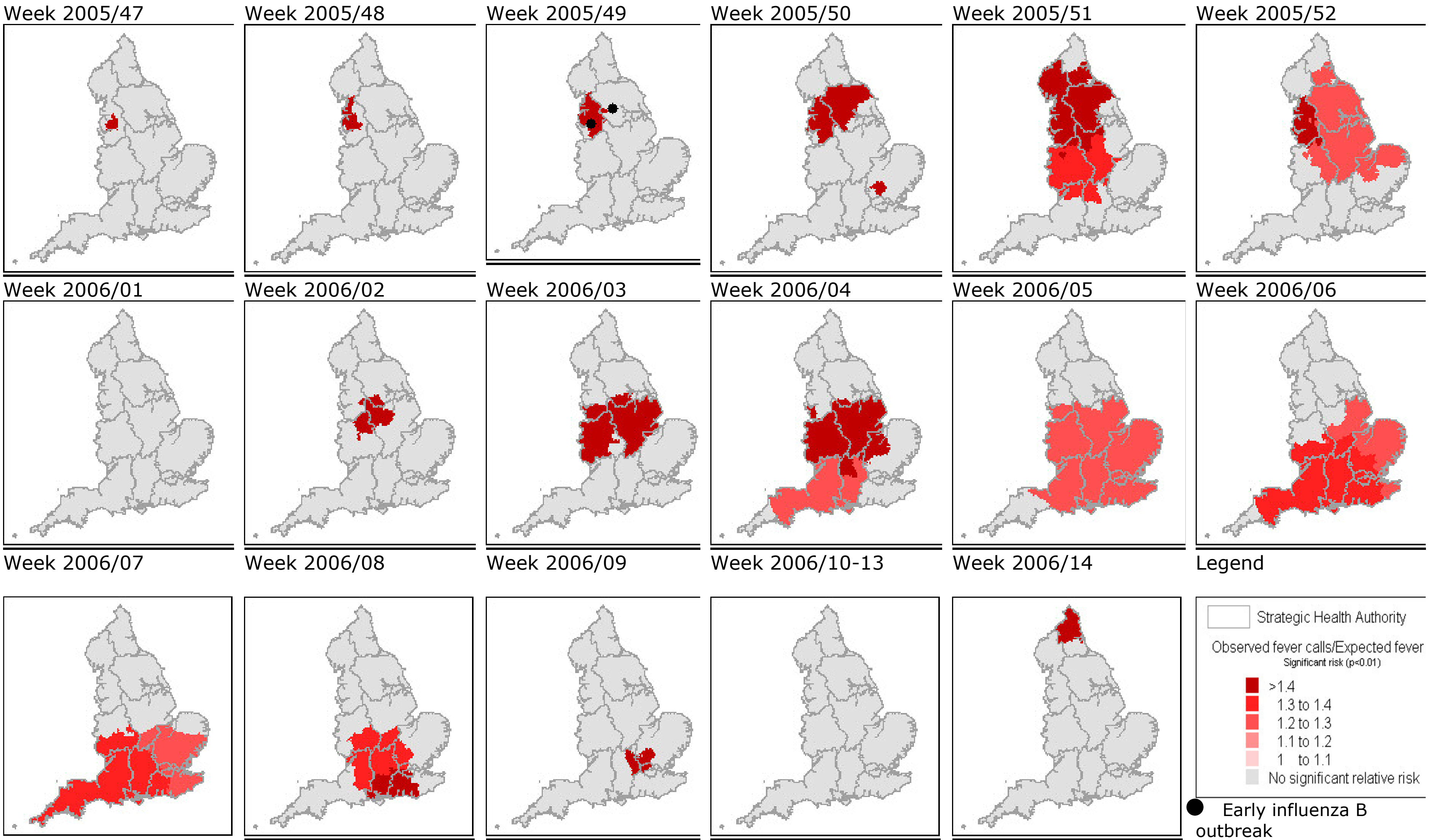
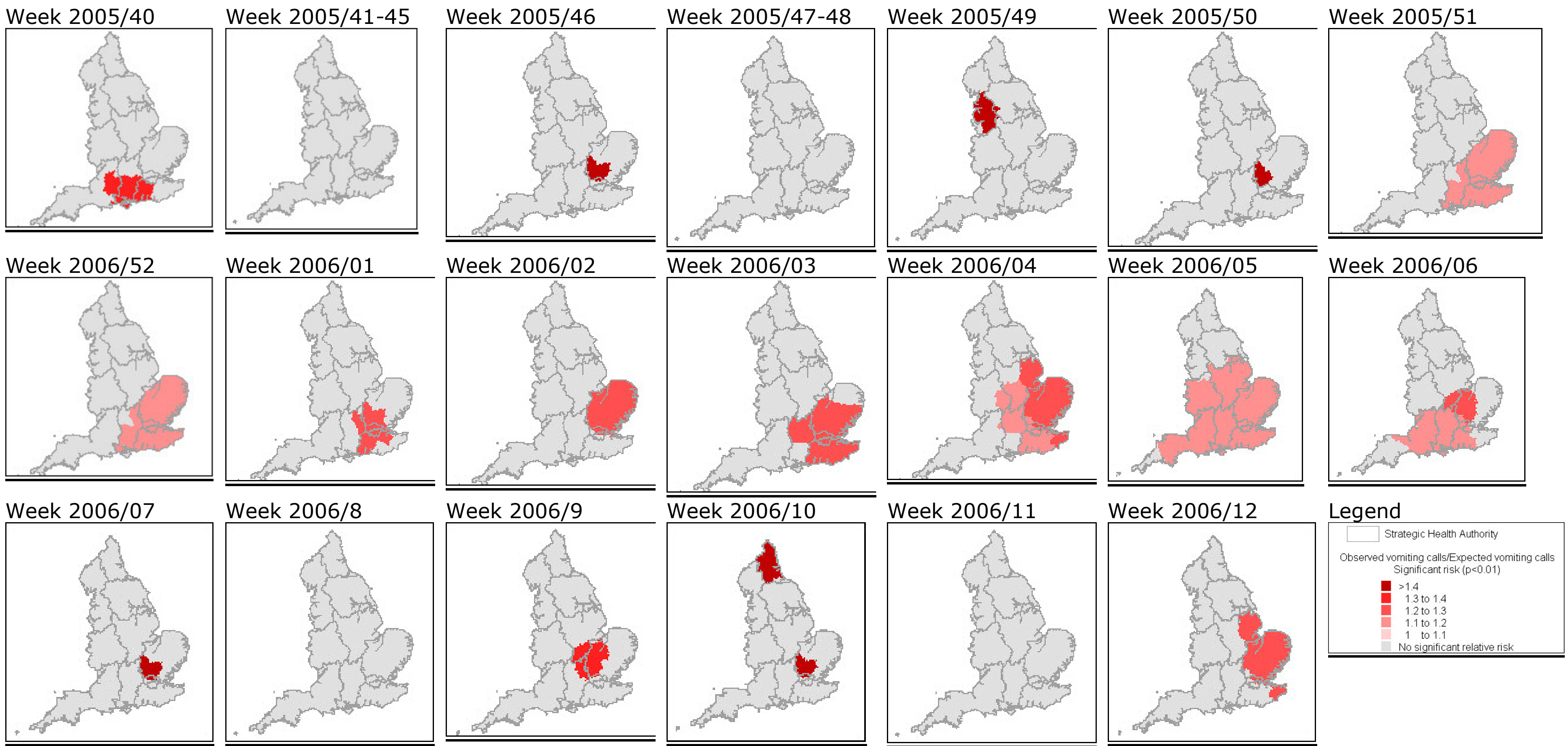


Figure 4. Areas with significantly high numbers of vomiting calls (clusters) displayed as observed/expected ratios by week. There were no significant clusters prior to 2005/40 and after 2006/12 during our test period.



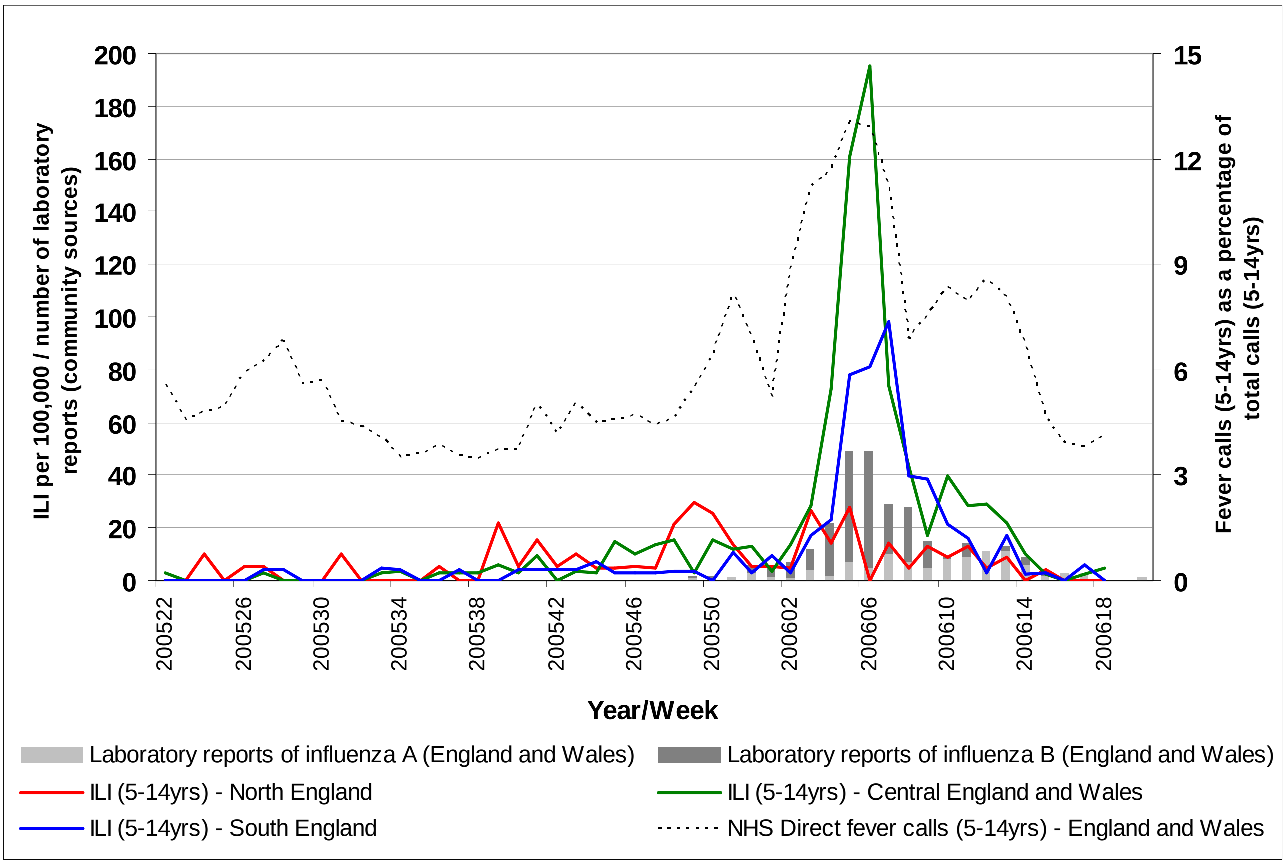
Spatio-temporal analysis of NHS Direct syndromic data should supplement rather than replace the current analyses of regional call data [16] and other components of the UK influenza surveillance programme. It may help to identify sub-regional variation in influenza rates that may not manifest themselves within data from sentinel networks of General Practitioners covering only a proportion of the population [46]. The added value, demonstrated by our retrospective analyses, was the early identification of a local rise in fever calls; a syndromic 'signal' that may require further investigation via conventional public health means. Used prospectively our analyses could be used to warn hospitals, doctors, schools and nursing homes of impending problems; identify areas to introduce or enhance microbiological sampling; and describe geographic variation in influenza incidence throughout an epidemic period. We used fever calls about school aged children because we were studying a national influenza B outbreak. Further spatial analyses of data about the NHS Direct 'colds and flu' syndrome are already used for national surveillance of influenza-like-illness in adults [16, 47]. Further winters' spatial analysis of respiratory syndromes is recommended to assess operational reliability of these results, and also to monitor influenza A strains associated with a high adult incidence.

Conclusions

We have retrospectively described how NHS Direct syndromic call data can be used for geographical surveillance of influenza and norovirus infection in England. Spatio-temporal analyses of fever calls for the 5-14yr age group provided a timely and unique description of the evolution of a national influenza outbreak. These data provide the opportunity to present how a virus such as influenza spreads across the Country affecting different localities to different degrees during the course of the influenza season. Such a tool therefore has utility for local services for both seasonal influenza and in the event of an influenza pandemic. In a similar way the tool may be useful for tracking norovirus, though lack of reliable and consistent comparison data make this more difficult to assess. In interpreting these results, care must be taken to consider other infectious and non-infectious causes of fever and vomiting.

To our knowledge this is the first published spatio-temporal analysis of national telehealth data used for surveillance purposes. The scan statistic should be considered analyses of telehealth data elsewhere. We will use these results to initiate prospective geographical surveillance of influenza in England.

Figure 5. Weekly GP influenza-like-illness consultation rates (ILI) for the 5-14yr age group for North, Central and Southern England; weekly numbers (all ages combined) of positive influenza samples (influenza A and B separately) from community sources; and NHS Direct fever calls as a percentage of total calls (5-14yrs) for England and Wales.



References

1. Henning KJ: What is syndromic surveillance? *MMWR Morb Mortal Wkly Rep* 2004, 53 Suppl:5-11.
2. Heffernan R, Mostashari F, Das D, Karpati A, Kulldorff M, Weiss D: Syndromic surveillance in public health practice, New York City. *Emerg Infect Dis* 2004, 10(5): 858-864.
3. Steiner-Sichel L, Greenko J, Heffernan R, Layton M, Weiss D: Field investigations of emergency department syndromic surveillance signals - New York City. *MMWR* 2004; 53 Suppl:184-9.
4. Goddard NL, Kyncl J, Watson JM: Appropriateness of thresholds currently used to describe influenza activity in England. *Commun Dis Public Health* 2003 Sep; 6(3):238-45.
5. Tompkins DS, Hudson MJ, Smith HR, et al: A study of infectious intestinal disease in England: microbiological findings in cases and controls. *Commun Dis Public Health* 1999, 2(2): 108-113.
6. Lopman BA Lopman BA, Adak GK, Reacher MH, Brown DW. et al: Two Epidemiologic Patterns of Norovirus Outbreaks: Surveillance in England and Wales, 1992-2000. *Emerging Infectious Diseases* 2003*;* 9(1): 71-77.
7. Desai S, Zhao H, Cooke MK, et al: Surveillance of Influenza and other respiratory viruses in the United Kingdom: October 2005 to May 2006. *Commun Dis Rep Suppl* 2006, 16(39): 1-6.
8. Johnston CP, Qiu H, Ticehurst JR, Dickson C, Rosenbaum P, Lawson P, Stokes AB, Lowenstein CJ, Kaminsky M, Cosgrove SE, Green KY, Perl TM: Outbreak management and implications of a nosocomial norovirus outbreak. *Clin Infect Dis* 2007 Sep 1;45(5):534-40. Epub 2007 Jul 18.
9. Akazawa M, Sindelar JL, Paltiel AD: Economic costs of influenza-related work absenteeism. *Value Health* 2003, 6(2):107-15.
10. Cliff AD, Haggett p, Ord KL, Versey G: Spatial diffusion: An historical Geography of Epidemics in an Island Community 1981. Cambridge University Press. Chapter 2. The study of spatial diffusion.
11. Smallman-Raynor M, Johnson N, Cliff A: The spatial anatomy of an epidemic: influenza in London and the county boroughs of England and Wales, 1918-1919. *Trans Inst Br Geogr* 2002, 452-470.
12. Cooke MK, Crofts JP, Joseph CA, et al: Influenza and other respiratory viruses survellance in the United Kingdom October 2003 to May 2004. *Commun Dis Rep Suppl* 2005; 15(1): 1-8.
13. Influenza Surveillance in the United Kingdom. HPA Website. http://www.hpa.org.uk/infections/topics_az/influenza/seasonal/uk_data_sources.htm.
14. Wheeler JG, Sethi D, Cowden JM, et al: Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. *BMJ* 1999 Apr 17, 318(7190):1046-50.
15. Lopman B, Vennema H, Kohill E. et al: Increase in viral gastroenteritis outbreaks in Europe and epidemic spread of new norovirus variant. *Lancet* 2004, 363(9410):682-8.
16. Cooper DL. Case study: The Use of Tele-health Data for Syndromic surveillance in England and Wales. In: Lombardo JS, Buckeridge DL, eds. *Disease Surveillance: A Public Health Informatics Approach.* Wiley: New Jersey 2007; Chap 8: 335-365.
17. Smith G, Cooper D, Loveridge P, Chinemana F, Gerard E, Verlander N: A national syndromic surveillance system for England and Wales using calls to a telephone helpline. *Euro Surveill 2006*, 11(12):220-4.
18. Loveridge P, Cooper DL, Harris J, Grey J, Smith GE. Comparison of Trends in NHS Direct Gastroenteric Syndromes with Norovirus Outbreaks & Laboratory Data 2004 – 2007. Proceedings Health Protection 2007.
19. Kulldorff M: A spatial scan statistic. Bernoulli and Poisson Models. Communications in Statistics: Theory and Methods, 26:1481-1496, 1997.
20. Boscoe FP, McLaughlin C, Schymura MJ, Kielb CL: Visualization of the spatial scan statistic using nested circles. *Health and Place* 2003; 9(3):273-277.
21. Kulldorff M, Heffernan R, Hartman J, Assuncao R, Mostashari F: A space-time permutation scan statistic for disease outbreak detection. *PLoS Med* 2005 Mar;2(3):e59. Epub 2005 Feb 15.
22. Jones RC, Liberatore M, Fernandez JR, Gerber SI: Use of a prospective space-time scan statistic to prioritize shigellosis case investigations in an urban jurisdiction. *Public Health Rep* 2006;121(2):133-9.
23. Mostashari F, Kulldorff M, Hartman JJ, Miller JR, Kulasekera V: Dead bird clustering: A potential early warning system for West Nile virus activity. *Emerging Infectious Diseases*x 2003, 9:641-646.
24. Pearl DL, Louie M, Chui L, Dore K, Grimsrud KM, Leedell D, Martin SW, Michel P, Svenson LW, McEwen SA: The use of outbreak information in the interpretation of clustering of reported cases of

Escherichia coli O157 in space and time in Alberta, Canada, 2000-2002. *Epidemiol Infect* 2006, 34(4):699-711.
25. Fleming DM: Weekly Returns Service of the Royal College of General Practitioners. *Commun Dis Public Health* 1999 Jun; 2(2):96-100.
26. O'Cathain A, Nicholl J, Sampson F, et al. Do different types of nurses give different triage decisions in NHS Direct? A mixed methods study. J Health Serv Res Policy 2004 Oct;9(4):226-33.
27. Knowles E, Munro J, O'cathain A, Nicholl J. Equity of access to health care. Evidence from NHS Direct in the UK. J Telemed Telecare 2006;12(5):262-5.
28. Salisbury C. The demand for out-of-hours care from GPs: a review. Fam Pract 2000;17(4):340-7.
29. Burt J, Hooper R, Jessopp L. The relationship between the use of NHS Direct and deprivation in southeast London: an ecological analysis. J Publ Hlth Med 2003; 25:174-176.
30. Rowlands S, Moser K. Consultation rates from the general practice research database. Br J Gen Pract 2002; 52:658-660.
31. St George, I. Cullen, M. Branney, M. The deprivation profile and ethnicity of Healthline callers. New Zealand Family Physician 2006 ; 33(6): 386-389.
32. Turner VF, Bentley PJ, Hodgson SA, et al. Telephone triage in Western Australia. Med J Aust 2002; 176(3):100-3.
33. Rolland E, Moore KM, Robinson VA, McGuinness D. Using Ontario's "Telehealth" health telephone helpline as an early-warning system: a study protocol. BMC Health Serv Res 2006;15;6:10.
34. Browne I, Cooper DL, Elam G, et al. Evaluation of the usefulness of syndromic surveillance – real time reassurance during major incidents. Health Protection 2007, Warwick University, September 2007.
35. Turcios RM, Curns AT, Holman RC, Pandya-Smith I, LaMonte A, Bresee JS, Glass RI; National Respiratory and Enteric Virus Surveillance System Collaborating Laboratories: Temporal and geographic trends of rotavirus activity in the United States, 1997-2004. *Pediatr Infect Dis J* 2006, 25(5):451-4.
36. Cooper DL, Smith GE, Edmunds WJ, Joseph C, Gerard E, George RC. The contribution of respiratory pathogens to the seasonality of NHS Direct calls. J Infect 2007; 55:240-248).
37. Openshaw S: The modifiable areal unit problem. *Concepts and Techniques in Modern Geography* 1984, 38: 41.
38. Armhein C: Searching for the elusive aggregation effect: Evidence from statistical simulations. *Environment & Planning* 1995, 27(1): 105-120.
39. Carrico R and Goss L: Syndromic surveillance: hospital emergency department participation during the Kentucky Derby festival. *Disaster Management & Response 2005*, 3(3): 73-79.
40. Sosin DM, DeThomasis J: Evaluation challenges for syndromic surveillance--making incremental progress. *MMWR* 2004, 24;53 Suppl:125-9.
41. Kleinman KP, Abrams A, Mandl K, Platt R: Simulation for assessing statistical methods of biologic terrorism surveillance. *MMWR* 2005, 54 Suppl:101-8.
42. Aamodt G, Samuelsen SO, Skrondal A: A simulation study of three methods for detecting disease clusters. *Int J Health Geogr* 2006, 5:15.
43. Carrat F, Valleron AJ: Epidemiologic mapping using the "kriging" method: application to an influenza-like illness epidemic in France. *Am J Epidemiol* 1992, 135(11):1293-300.
44. Sakai T, Suzuki H, Sasaki A, Saito R, Tanabe N, Taniguchi K: Geographic and temporal trends in influenzalike illness, Japan, 1992-1999. *Emerg Infect Dis* 2004, 10(10):1822-6.
45. Saito R, Paget J, Hitaka S, Sakai T, Sasaki A, van der Velden K, Suzuki H: Geographic mapping method shows potential for mapping influenza activity in Europe.*Euro Surveill* 2005, 10(10) :E051027.6.
46. Regan CM, Johnstone F, Joseph CA, Urwin M: Local surveillance of influenza in the United Kingdom: from sentinel general practices to sentinel cities? *Commun Dis Public Health* 2002, 5(1):17-22.
47. Cooper DL, Smith GE, Hollyoak VA, Joseph CA, Johnson L, Chaloner R: Use of NHS Direct calls for the Surveillance of influenza - a second year's experience. *Commun Dis Public Health* 2002, 5 (2): 132-136.

Part 3

Surveillance results focussing on respiratory and gastrointestinal infections

9 What can analysis of calls to NHS Direct tell us about the epidemiology of gastrointestinal infections in the community?

Published as: Cooper DL, Smith GE, O'Brien SJ, Hollyoak V, Long S. What can analysis of calls to NHS Direct Tell us about the Epidemiology of Gastrointestinal Infections in the community? *J Infect* 2003*; 46(2): 101-105.*

Abstract

Objectives: Most gastrointestinal (GI) illness within the United Kingdom goes undetected by routine surveillance. A national telephone helpline for health advice (NHS Direct) offers a new source of data on GI infection. We aim to describe NHS Direct calls suggestive of GI infection and the outcome of these calls.

Methods: Details of over 150,000 telephone calls were collected from NHS Direct over a 6 month period. Calls about ‘diarrhoea’, ‘vomiting’ or ‘food poisoning’ were defined as GI calls and described according to the age of the patient and call outcome.

Results: Gastrointestinal calls accounted for 10.3% of total calls (‘diarrhoea’ = 4.9%, ‘vomiting’ = 5.1%). GI calls as a proportion of total calls were significantly high in children under 1 year (23.5%) and aged 1-4 years (21.5%). Call outcomes which resulted in further NHS care being recommended accounted for 72.3% of total calls and 54.5% of GI calls.

Conclusions: A high proportion of NHS Direct calls were about GI symptoms especially for children under 5 years. When compared with all NHS Direct calls, GI calls were less likely to result in further NHS care being recommended. Analysis of NHS Direct calls provides further insight into GI infection in the community.

Introduction

At present, data from National Health Service (NHS) and Public Health Laboratory Service (PHLS) laboratories (laboratory confirmed cases), the Royal College of General Practitioners Weekly Returns Service (consultations for intestinal infection) and statutory notifications of food poisoning are used in the routine surveillance of GI infection in England and Wales [1-2]. However, most GI infection within the community in the UK is undetected by surveillance [3-4], prompting calls for measures to supplement these routine data sources [5].

NHS Direct is a nurse led national telephone helpline providing health advice in the UK. The service is open 24 hours a day, 365 days of the year and responds to over 5 million calls a year. Nurses use a computerised clinical decision support system to handle calls (the NHS Clinical Assessment System (NHS CAS)). The NHS CAS uses approximately 200 computerised symptom based clinical algorithms. On receipt of a call and description of the caller's symptoms, nurses use their judgement to select the most appropriate algorithm. The nature and severity of the symptoms of the person about whom the call is being made, dictate a distinct line of questioning within an algorithm. NHS Direct algorithms form a tree-like structure of questions, with many possible pathways and several 'end points' (outcomes / dispositions). A disposition consists of guidance notes for 'self care' for the problem or referral of the caller to NHS primary or secondary care (e.g. General practitioners' (GP) surgery or Accident & Emergency (A&E) department). Although there are over 30 different disposition types, more than 85% of total calls result in one of 7 dispositions.

NHS Direct data have already been used for the surveillance of 'influenza-like illness' in England [6-7]. However, NHS Direct receives calls from patients reporting a wide range of symptoms, including those which may be suggestive of GI infection (e.g. vomiting and diarrhoea). This provides an opportunity to explore the data for the surveillance of GI infection, some of which may not be detected by routine surveillance schemes. We therefore sought to describe calls suggestive of GI infection and the outcome of these calls.

Materials and Methods

NHS Direct sites handle health information calls (e.g. requests for information on local health services) as well as calls relating to symptoms (‘triage calls’). Data on all ‘triage’ calls (from this point on referred to as ‘total calls’) were obtained from 3 NHS Direct sites (Midland Shires, West Midlands and West Yorkshire NHS Direct) for a 6 month period (May to November, 2001). The population covered by Midland Shires NHS Direct is approximately 3 million, West Midlands NHS Direct 2.3 million and West Yorkshire NHS Direct 2.1 million. Some NHS Direct sites also receive a proportion of their calls via local GP Co-operatives. In order to be consistent across sites, GP Co-operative calls were excluded from the analysis.

We selected ‘vomiting’, ‘diarrhoea’ and ‘food poisoning’ algorithms as those suggestive of GI infection. Calls for these algorithms were added together to calculate ‘GI calls’ and broken down by the following age groups; <1, 1-4, 5-14, 15-64 and >65 years. We categorised outcomes for all calls into one of 7 disposition groups (Table I).

Population based weekly call rates per 100,000 for total calls and GI calls were calculated for all age groups, using the Office for National Statistics mid term population estimates (2000).

Because of the fluctuations in NHS Direct site catchment populations (e.g. due to commuting) and sudden surges in demand due to advertising campaigns or out of hours calls (e.g. weekends, Bank Holidays), GI calls were also calculated as proportions of total calls. For each age group, algorithm and disposition group, the proportion of total calls was calculated with the 99% confidence intervals for these proportions.

Table I - NHS Direct call disposition (outcome) groups.

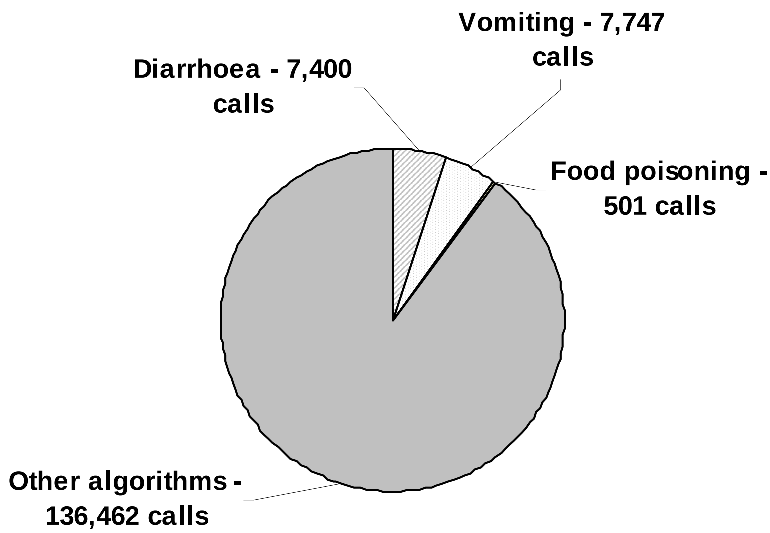
Disposition group	**Summary Description**
Home Care	*Appropriate advice on self help is provided by the nurse.*
GP Routine	*Assessment by a GP or dentist via a routine appointment (usually within 1 week).*
GP Urgent	*Assessment by a GP or dentist is recommended for the same day.*
A&E	*Assessment at casualty is needed urgently or within 4 hours.*
999	*Paramedic intervention is required.*
Community	*For situations where a community service is recommended (e.g. family planning clinic, social worker or mental health care team).*
Other	*e.g. NHS Direct nurse calls the poisons centre.*

Results

During the six month period the three NHS Direct sites responded to 152,110 total calls (West Yorkshire = 58,959, West Midlands = 51,821, Midland Shires = 41,330) of which 7,400 (4.9%) were about ‘diarrhoea’, 7,747 (5.1%) about ‘vomiting’ and 501 (0.3%) about self reported food poisoning (Figure1). All GI calls (15648) accounted for 10.3% of total calls.

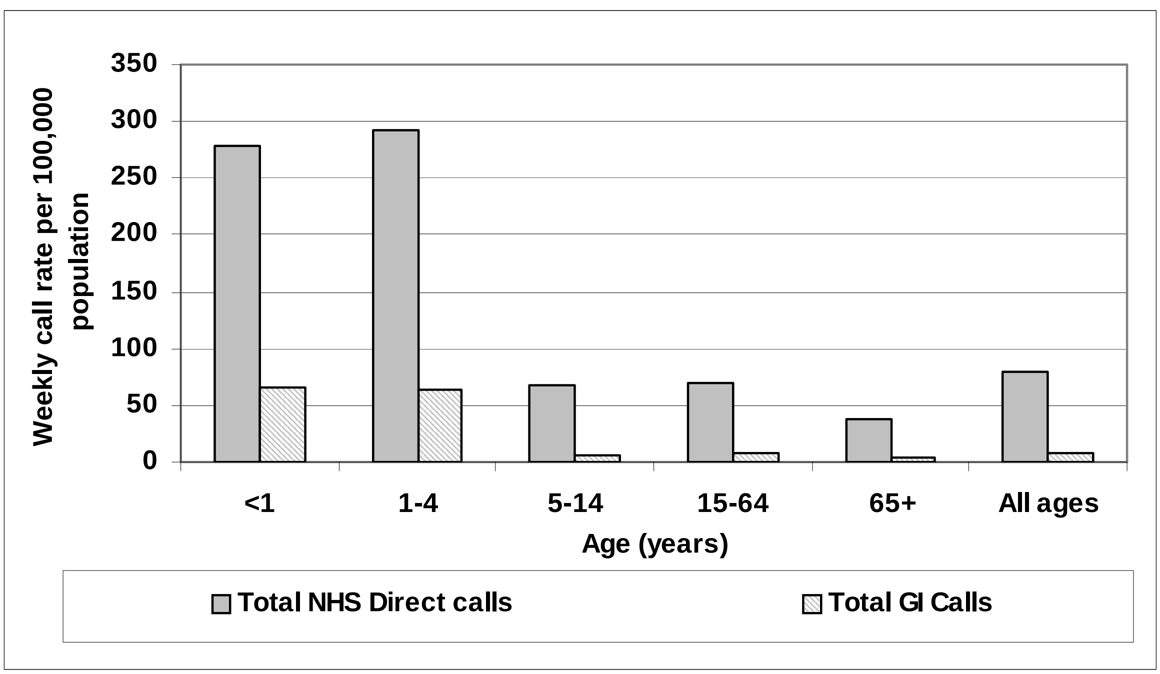
The GI call rate for all ages was 8 per 100,000 population per week. Rates were highest for children under 1 year (65 per 100,000) and for those aged 1-4 years (63 per 100,000) (Figure 2). This reflects age-group specific total call rates which were also high for the under 1 year (278 per 100,000) and 1 to 4 year age groups (293 per 100,000). The total call rate (all ages) was 79 per 100,000 population per week.

Figure 1 – NHS Direct total calls categorised by algorithm (N=152,110).



The proportion of GI calls made about children under 1 year (10.2% (99%CI 9.6-10.9)) and about those between 1 and 4 years (39.7% (38.7-40.7)) were significantly higher than the proportion of total calls made about children under 1 year (4.5% (4.3-4.6)) and about those between 1-4 years (18.9% (18.7-19.2)).

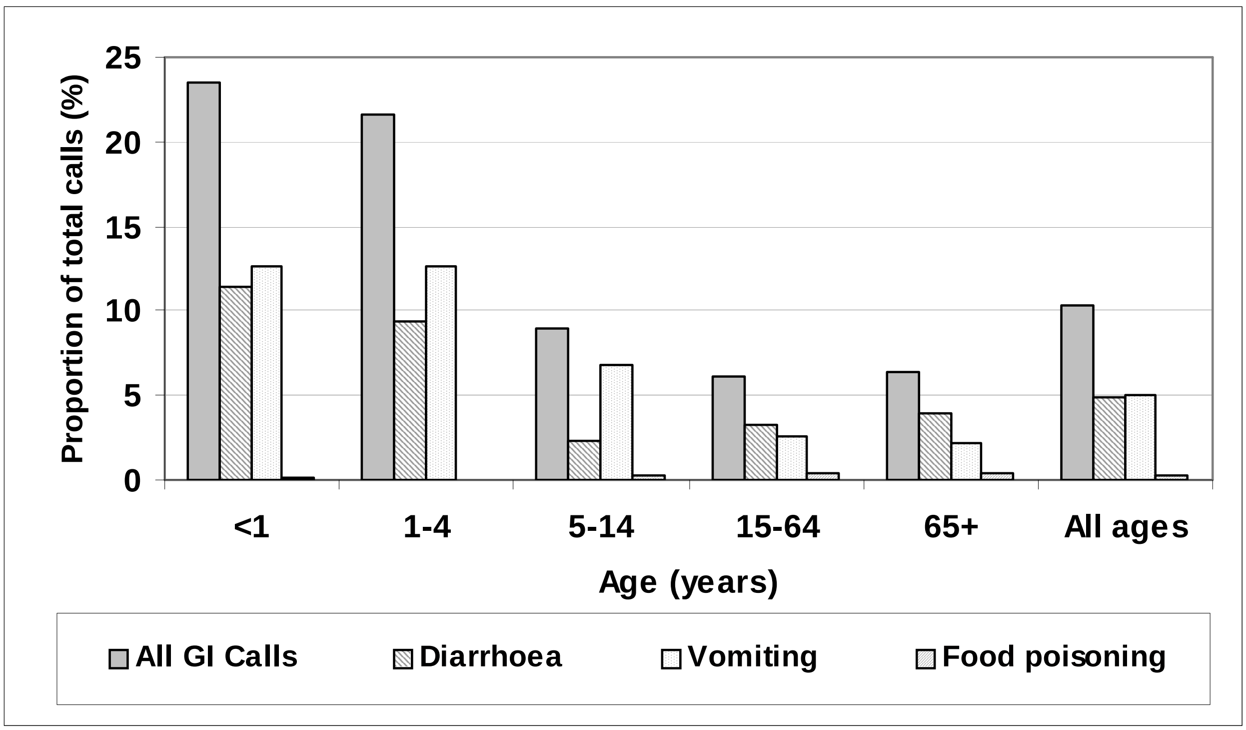
Figure 2 – Age group specific call rates for NHS Direct gastrointestinal and total calls.



Gastrointestinal calls as a proportion of total calls were significantly higher in children under 1 year (23.5% (99%CI 22.2-24.8)) and for those aged 1-4 years (21.5% (20.9-22.2)) when compared with all ages (10.3% (10.1-10.5)) (Figure 3).

The under 1 and 1-4 year age groups had the highest proportions of 'vomiting' and 'diarrhoea' calls (Figure 3). The lowest proportion of 'diarrhoea' calls was in the 5-14 year age group (2.3%) whilst the lowest proportion of 'vomiting' calls was in the over 65 years age group (2.1%). The proportion of total calls about food poisoning was negligible (0.3%).

Figure 3 – Gastrointestinal calls as a proportion of total calls by age group and algorithm.



The proportion of calls resulting in a 'home care' disposition was significantly higher for GI calls (41.5% (40.5-42.5)) (Figure 4) than for total calls (23.4% (23.1-23.7) (Figure 5) in all age groups. The proportions of GI calls resulting in a 'GP Urgent' disposition were highest for children. In the 5-15 year age group the proportion of GI calls resulting in a 'GP Urgent' disposition was significantly higher than for any other age group (48.5% (45.2-51.7)).

Figure 4 – Proportion of gastrointestinal calls by disposition (outcome) group and age group.

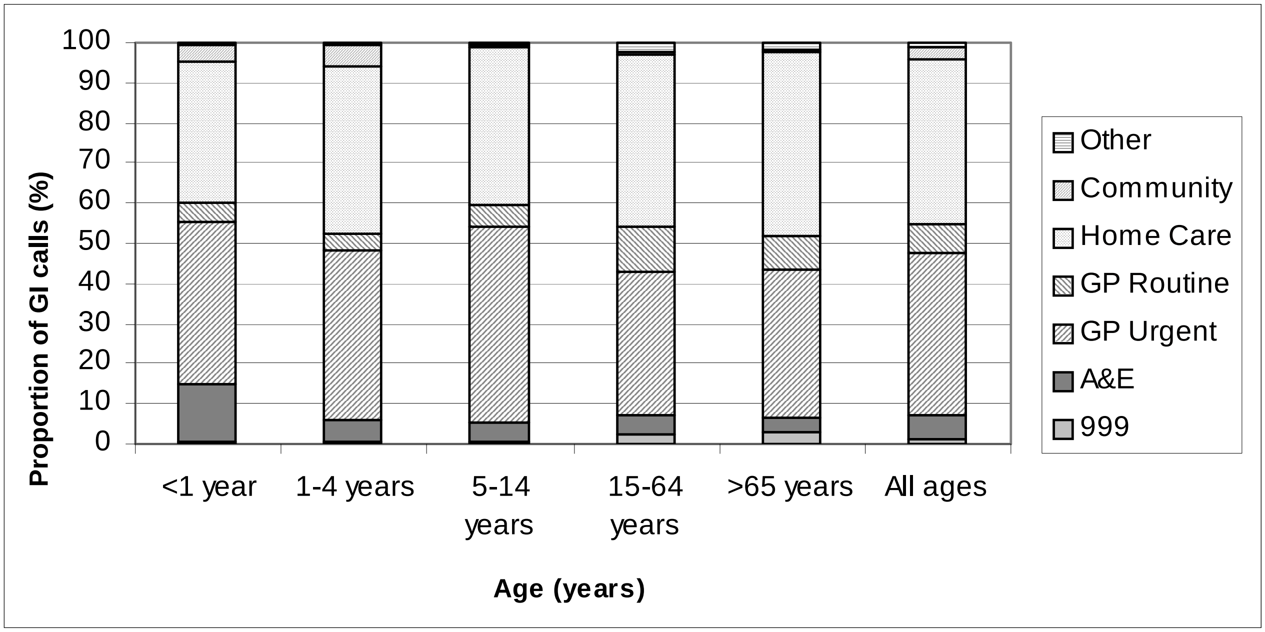
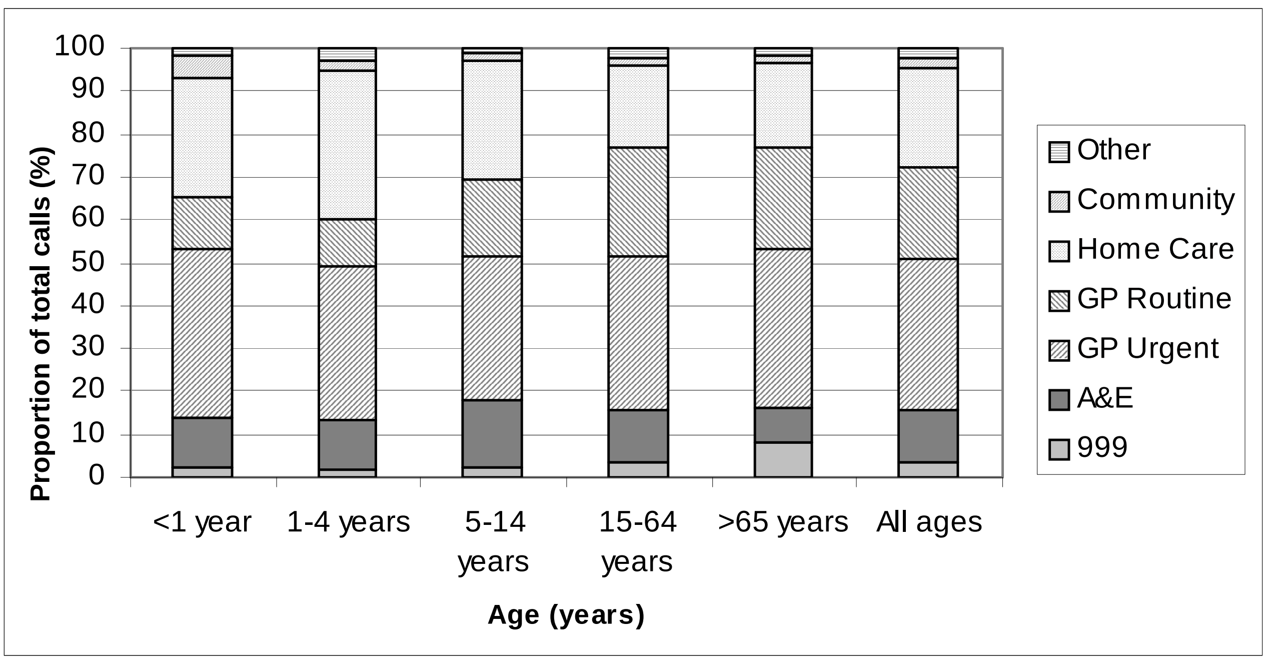


Figure 5 – Proportion of total calls by disposition (outcome) group and age group.



The proportion of calls resulting in a '999' disposition was significantly lower for GI calls (1.2% (1-1.4)) than total calls (3.5% (3.3-3.6)) in all age groups. Similarly, the proportion of calls resulting in an 'A&E' disposition was also significantly lower for GI calls (6.1% (5.6-6.6)) than total calls (12.1% (11.9-12.3)) in all age groups of 1 year or older. In the under 1 year age group the proportion of GI calls which result in an ‘A&E’ disposition was 14%, more than double that of any other age group (Figure 4).

Discussion

Ten percent of calls to NHS Direct are about GI symptoms. Assuming a national NHS Direct call rate equal to that of the three sites examined, this approximates to over 200,000 calls a year about GI symptoms. This number is likely to increase to reflect the observed year on year increases in NHS Direct call volumes since the service achieved national coverage in 2000.

Calls about children aged under five years account for about a third of all NHS Direct calls, whereas calls about GI symptoms in the under five year olds account for approximately half of all GI calls. This further demonstrates high usage of the service by parents of young children which is described elsewhere [8], particularly by those whose children have gastrointestinal symptoms.

‘Vomiting’ is reported more frequently for children and ‘diarrhoea’ more frequently by adult callers to NHS Direct. This may partly be explained by the age distribution and clinical presentation of common pathogens. For example, rotavirus and Norwalk-like viruses commonly cause vomiting illness in children. Unpublished NHS Direct data indicate that the highest proportion of ‘vomiting’ calls made about children occurs during January to March and coincides with a peak in laboratory detections of rotavirus and Norwalk-like virus. As more call data about GI symptoms become available, the effect of seasonality of GI infection on the distribution of GI calls will be further investigated.

Low numbers of ‘food poisoning’ calls (0.3% of total calls) are described. A food-borne aetiology for a caller’s symptoms is difficult to establish from a brief telephone interview. NHS Direct nurses are trained to use symptom based algorithms where possible (e.g. vomiting) rather than algorithms that by their titles suggest a diagnosis such as ‘food poisoning’. Food poisoning is unique amongst the list of notifiable diseases in that it requires the patient (or their clinician) to ascribe a cause to their symptoms. Based on symptoms alone an individual cannot be certain if their symptoms were due to foodborne, person to person or environmental transmission, with the possible exception of those involved in recognised and proven point source foodborne outbreaks. Some individuals with gastrointestinal illness will improperly blame contaminated food [9]. The usefulness of the ‘food poisoning’ algorithm as an indicator of GI disease is therefore questionable and it is likely that this algorithm will be excluded from future versions of the NHS CAS.

Calls about GI symptoms, compared with all calls (total calls) to NHS Direct, are less likely to result in further NHS care being recommended. This probably reflects the self-limiting nature of much GI infection. It may also demonstrate how the use of the NHS CAS can differentiate mild from more serious GI disease.

Although data are available from NHS Direct for individual algorithms, linkage of clusters of symptoms to individual callers is not possible. For example, whilst the ‘vomiting’ algorithm may be used to handle a call, diarrhoea and other symptoms may also be reported to the nurse as part of the questioning within this algorithm. Further development of this work to link symptoms to callers will be required to provide an indication of the severity of callers’ illnesses and possible clues to a causal pathogen (e.g. bloody diarrhoea associated with *Escherichia coli 0157* infection).

In conclusion, analysis of NHS Direct call data from three sites has demonstrated that a tenth of calls to NHS Direct are about symptoms suggestive of GI infection, half of these calls being about children. Nearly half of the callers with GI symptoms are not advised to consult other primary or secondary care services and are thus unlikely to be included in existing routine GI infection surveillance systems. This preliminary work has established useful baselines and will be extended to collect data from all NHS Direct sites in England and Wales. NHS Direct data collected in the future will add to the knowledge about the burden of gastrointestinal disease in England and Wales and be of use in monitoring the Food Standards Agency target to reduce the burden of foodborne infectious disease by 20% by 2006 [9].

References

1. Wall PG, Louvois J, Gilbert RJ, et al. Food poisoning: notifications, laboratory reporting, and outbreaks: where do the statistics come from and what do they mean? *CDR Review* 1996;**6**:93-100.
2. Fleming D. Weekly Returns Service of the Royal College of General Practitioners. *Commun Dis Pub Health* 1999;**2**:96-100.
3. Foods Standards Agency. A report of the Study of Infectious Intestinal Disease in England. *The Stationery Office* 2000. London.
4. Palmer S, Houston H, Lervy B, et al. Problems in the diagnosis of foodborne infection in general practice. *Epidemiol Infect* 1996;**117**(3):479-84.
5. Wheeler GW, Sethi D, Cowden JM, et al. Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. *BMJ* 2000;**318**:1046-50.
6. Harcourt S, Smith GE, Hollyoak V, et al. Can calls to NHS Direct be used for Syndromic Surveillance? A pilot at three sites using influenza as an example. *Commun Dis Public Health*. 2001;**4**:178-82.
7. Cooper DL, Smith GE, Hollyoak VA, et al. Use of NHS Direct calls for the surveillance of influenza – a second year's experience. *Commun Dis Public Health 2002;***5**(2):132-6.
8. Payne F, Jessopp L. NHS Direct: review of activity data for the first year of operation at one site. *J Pub He Med* 2001;**23**:155-158.
9. Adak GK, Long SM, O'Brien SJ. Trends in Indigenous Foodborne Disease (IFD) and Deaths, England and Wales – 1992 to 2000. *Gut* (in press).
10. Food Standards Agency: Strategic Plan 2001-2005. FSA/0020/0501. 2001 < http://www.foodstandards.gov.uk/multimedia/pdfs/stratplan.pdf>

10 Use of NHS Direct calls for the surveillance of influenza – a second year's experience

Published as: Cooper DL, Smith GE, Hollyoak VA, Joseph CA, Johnson L, Chaloner R. Use of NHS Direct calls for the surveillance of influenza – a second year's experience. *Commun Dis Public Health 2002; 5(2): 132-136.*

Abstract

A second year's call data to NHS Direct are presented as an evaluation of their usefulness for influenza surveillance.

During the winter of 2000/01, age-group specific data relating to the 'cold/flu' algorithm were collected from 6 NHS Direct sites (population coverage: 16 million). Numbers of total calls were collected from all 23 NHS Direct sites on a daily basis. Despite the winter of 2000/01 having been a season of low activity for influenza in the United Kingdom, NHS Direct data demonstrated a peak in the 'cold/flu' calls as a proportion of the total calls (3.1% (672 'cold/flu' calls) during week 06/01). This coincided with the peak recorded by routine influenza surveillance systems. In addition there was an earlier peak in the proportion of 'cold/flu' calls (3.3% during weeks 52/00 (789 'cold/flu' calls) and 01/01 (749 'cold/flu' calls)) which may have been due to other respiratory infections, the lack of specificity of the definition of NHS Direct 'cold/flu' calls and increased 'out of hours' calls to NHS Direct.

Despite limitations, the timeliness of NHS Direct data, the total population coverage of the service and the ability to provide local information on 'cold/flu' calls make the call data suitable for further surveillance during the winter of 2001/2002. It is hoped that as NHS Direct reaches a 'steady state', in terms of population coverage and uniformity of clinical support systems, it will be possible to begin to construct 'baselines' for the respiratory disease related call data.

Introduction

Commitment from the NHS to provide easy access to healthcare services has led to the implementation of NHS Direct; a nurse led national telephone helpline for health advice, open 24 hours, 365 days of the year. NHS Direct nurses use decision support software to guide the advice given in response to a call. Diagnoses are not made. The 'real time' availability of NHS Direct data on approximately 100,000 calls a week has created new possibilities for disease surveillance. Although there has been some work examining NHS Direct call patterns [1-3], there are few examples of NHS Direct data, or health helpline call data from other countries, being used for communicable disease surveillance [4-5]. One example from Milwaukee showed how analysis of calls to a nurse helpline could track an outbreak of cryptosporidiosis [6]. This analysis was, however, performed on data derived from calls about an extreme event (50% of the City's population were infected), and it remains to be seen whether surveillance using data derived from helplines can detect less dramatic outbreaks.

Routine surveillance of influenza in England and Wales is performed by the Public Health Laboratory Service (PHLS) Enteric, Respiratory and Neurological Virus Laboratory (ERNVL), PHLS Communicable Disease Surveillance Centre (CDSC) and the Birmingham Research Unit of the Royal College of General Practitioners (RCGP) [7-8] in collaboration with other NHS colleagues. NHS Direct data were first evaluated for their potential use for influenza surveillance during the winter of 1999/2000 when three different call handling systems were in operation and NHS Direct covered only part of the country. By winter 2000/01, NHS Direct covered all of England and Wales and used the same call handling system at 9 of its 23 sites, providing a good opportunity to extend its evaluation as a means of influenza surveillance into a second winter.

The objectives of the extended evaluation were:

- to compare data from NHS Direct calls with data from routinely used surveillance systems for influenza;

- to compare two years' NHS Direct data and evaluate the usefulness and timeliness of the data as indicators of influenza activity.

Methods

Data collection

During winter 2000/01, we received data on the daily total of calls to all 23 NHS Direct sites. The population covered by each NHS Direct site was estimated from the Office for National Statistics mid term population estimates and used to calculate NHS Direct total call rates. NHS Direct total call rates were compared with those from 1999/2000.

Six of the 9 NHS Direct sites which used a single clinical decision support system, the NHS Clinical Assessment System (CAS), also provided numbers of NHS Direct 'cold/flu' calls by the following age groups (years): 0-4,5-14,15-44,45-64 and >65 years. The age recorded was that of the person about whom the call was made. A 'cold/flu' call was defined as a call where the caller's description of their symptoms led the NHS Direct nurse to use a 'cold/flu' clinical algorithm (note diagnosis are not made during the handling of a call). The data on 'cold/flu' calls for the previous week (Monday-Sunday) were usually provided on a Monday morning from four sites (Manchester, North & Central London, North-East London and Kent-Surrey-Sussex) (Figure 1). Two sites provided 'cold/flu' data on a daily basis (Thames Valley and West Yorkshire).

The number of calls for 'cold/flu' as a proportion of all calls were compared with routine influenza surveillance data.

Data dissemination

NHS Direct data were summarised in a weekly bulletin and distributed by email to colleagues within the PHLS, NHS (including Regional Directors of Public Health) and NHS Direct.

Results

The total NHS Direct call rate peaked initially at 190 per 100,000 population (100,624 calls answered) during week 52/00 and then showed a second peak of 178 per 100,000 population (94,456 calls answered) during week 05/01. The total call rate increased from week 48/00 by 33% during the first peak and by 25% during the second peak. During the winter of 1999/2000, a peak in the NHS Direct total call rate occurred in week 52/99 (331 per 100,000) with an increase of 125% from week 48/99.

‘Cold/flu’ calls as a proportion of total calls peaked at 3.3% during weeks 52/00 (789 ‘cold/flu’ calls) and 01/01 (749 ‘cold/flu’ calls) before dropping to 2.3% (454 ‘cold/flu’ calls) in week 02/01 (Figure 2). A second peak in the proportion of ‘cold/flu’ calls (3.1% (672 ‘cold/flu’ calls)) occurred in week 06/01 and gradually reduced to 1.2% (265 ‘cold/flu’ calls) in week 12/01. The daily ‘cold/flu’ data from 2 sites are shown in Figure 3.

At the time of the first peak in winter 2000/2001 (week 52/00), the greatest increase in proportion of calls that were ‘cold/flu’ was for 15-44 year olds (Figure 4). During the second peak in the proportion of ‘cold/flu’ calls (week 06/01) the 5-14 year age group showed the largest percentage increase from week 48/00.

The second peak in the NHS Direct proportion of ‘cold/flu’ calls, the numbers of community based influenza isolates and the consultation rate for influenza and ‘influenza like illness’ all peaked during week 06/01 (Figure 5). The highest proportion of age specific NHS Direct ‘cold/flu’ calls, the highest age-specific RCGP consultation rate for influenza and ‘influenza like illness’ and the highest age specific proportion of community samples positive for influenza were all within the 5-14 year age group during winter 2000/2001 and peaked in weeks 04/01, 06/01 and 06/01 respectively.

Figure 1 – NHS Direct coverage of England and Wales

Map produced at the CDSC West Midlands - data subject to Crown Copyright

Figure 2 - Total NHS Direct call rates, winter 1999/2000 and winter 2000/2001, and proportion of calls classified as 'cold/flu' calls, winter 2000/2001.

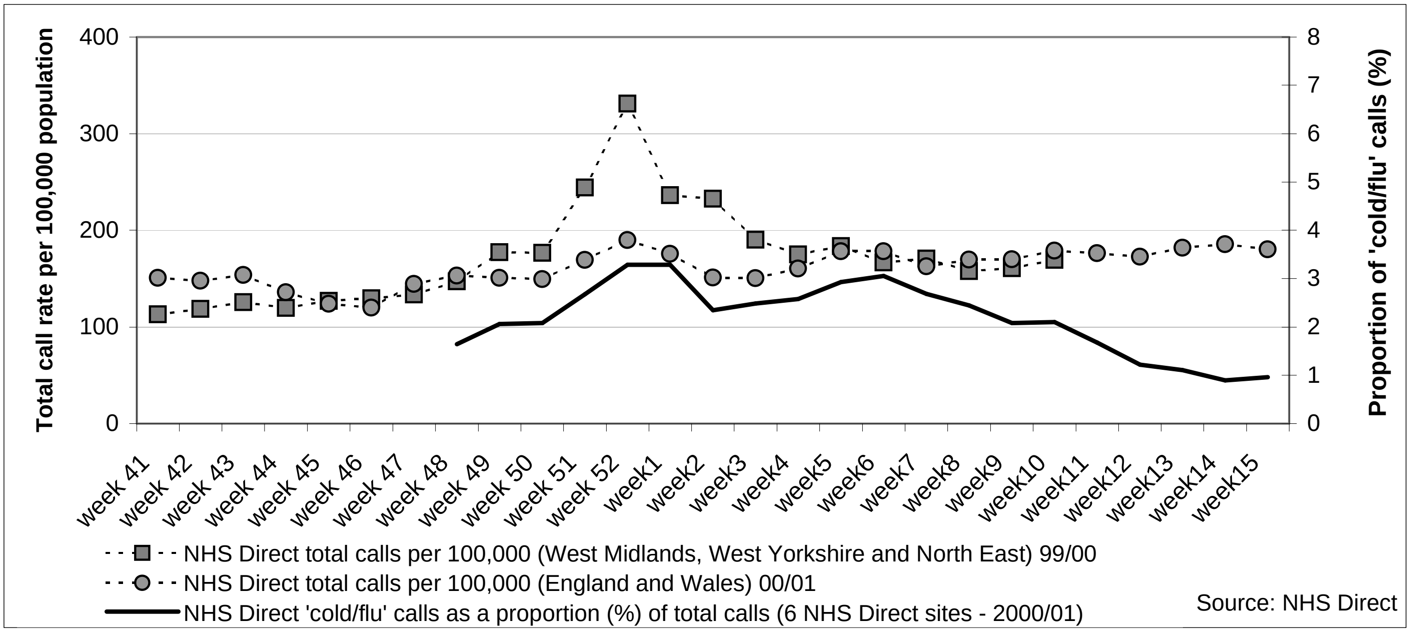


Figure 3 – 3 day rolling average of the proportion (%) of calls classified 'cold/flu' made to 2 NHS Direct sites, winter 2000/2001.

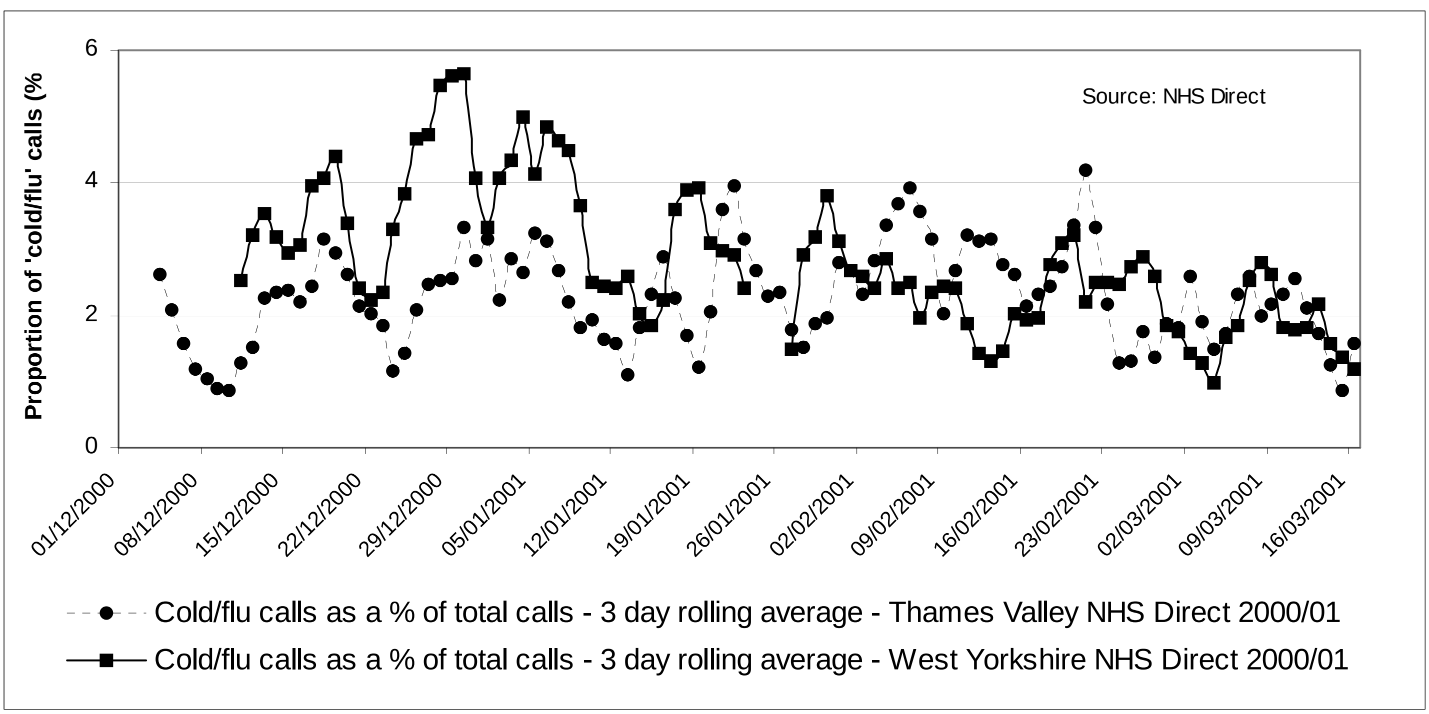


Figure 4 - Age specific percentage increase in the proportion of calls classified as 'cold/flu' for weeks 52/00 and 06/01 (from the baseline of week 48/00).

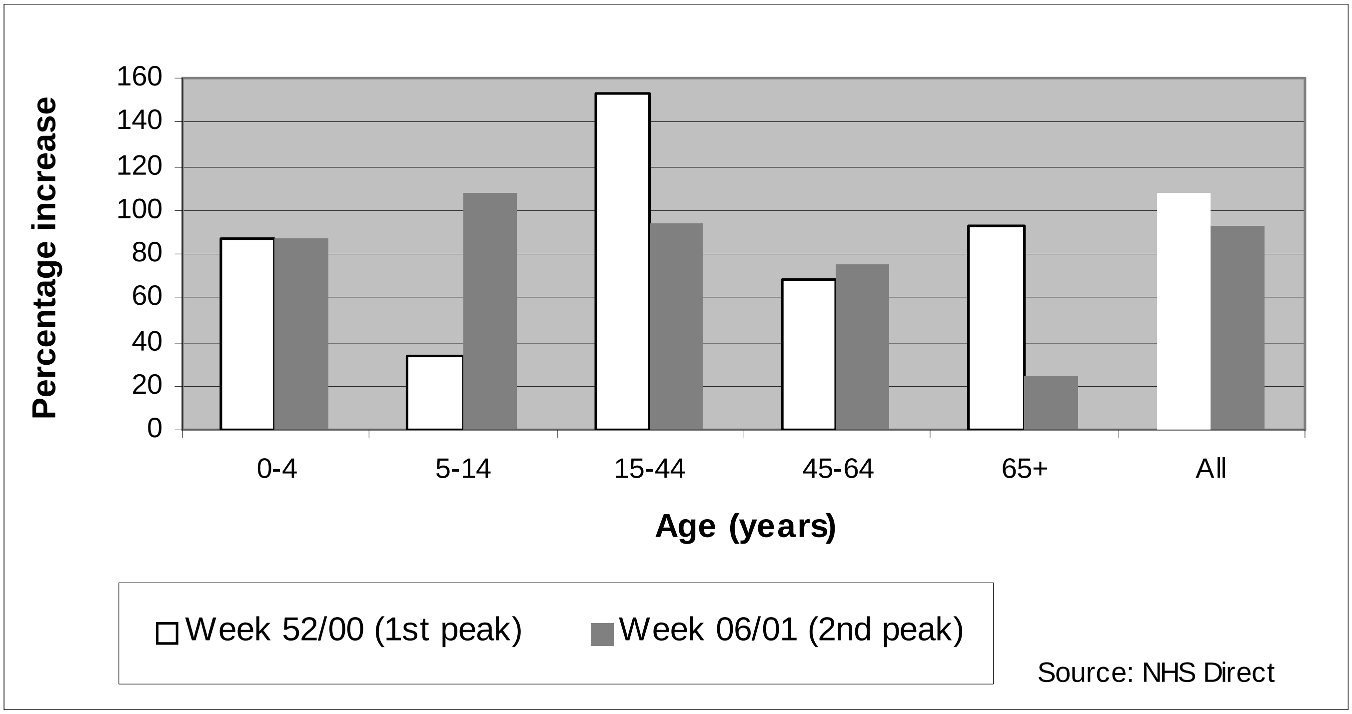
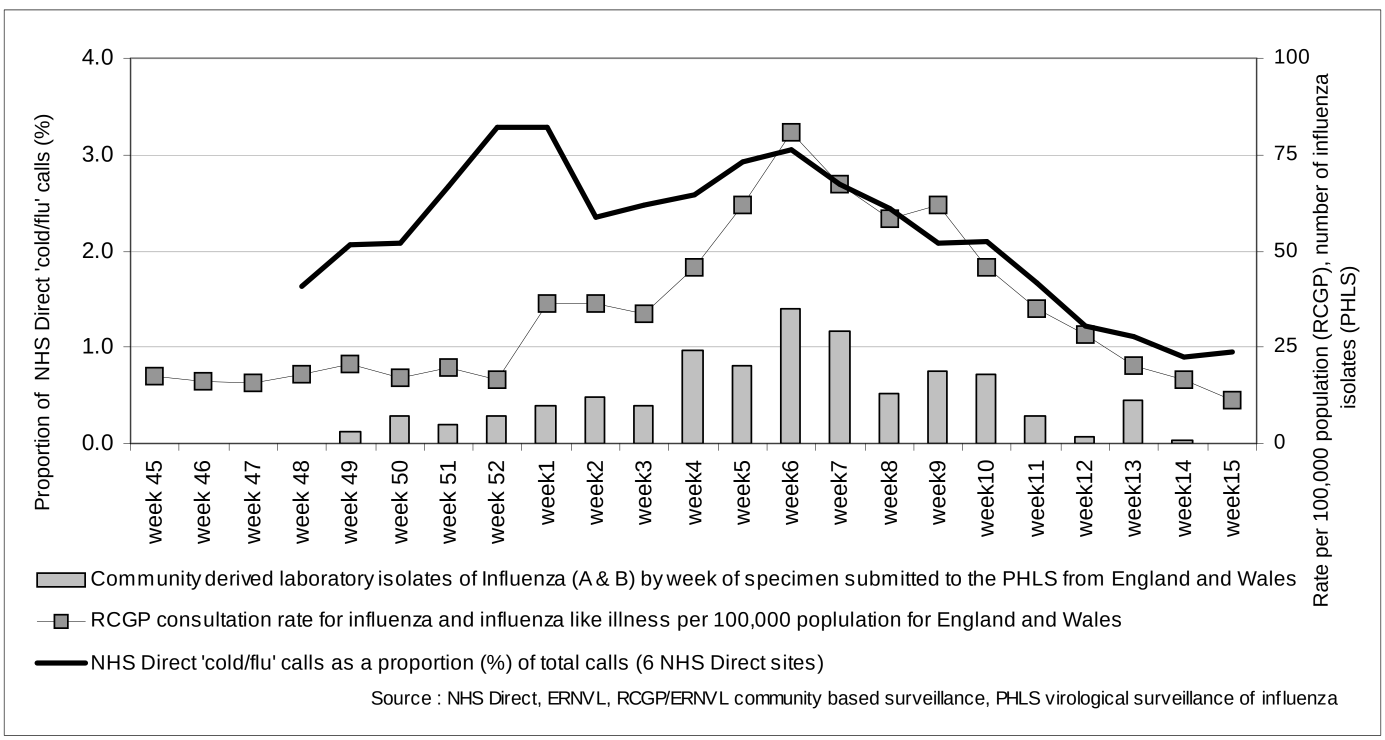


Figure 5 - Comparison of NHS Direct, RCGP and PHLS data – Influenza surveillance winter 2000/2001.



Discussion

Despite the winter of 2000/01 having been a season of low influenza activity in the United Kingdom (the peak RCGP consultation rate for influenza and 'influenza like illness' was the second lowest since1990), NHS Direct data demonstrated a peak in proportion of 'cold/flu' calls which coincided with the peak recorded by routine influenza surveillance systems.

There are several reasons for caution in the interpretation of NHS Direct call data:

- Between August and December 2000, 6 new NHS Direct sites were opened and 7 sites expanded. It takes months for calls to new sites to reach a stable rate and successive waves of advertising may affect this.
- Numbers of calls to NHS Direct call centres may vary in urban areas because of commuters making calls from within their work area rather than from their area of residence.
- During times of peak demand or staff training, calls are switched (networked) between the 23 NHS Direct sites.
- When compared to GP contact rates for all diseases, NHS Direct total call rates are low. The average contact rate for all disease in the Weekly Returns Service is 8,000 per 100,000 people per week, forty times greater than the NHS total call rate (200 per 100,000 per week) (A Ross personal communication).
- The lack of established baselines for NHS Direct call activity (compared with the long established baselines of the RCGP Weekly Returns Service).

The proportion, rather than the rate, of 'cold/flu' calls were used in the weekly bulletin to account for these effects.

Other than an increase in influenza activity, the Christmas peak in the proportion of 'cold/flu' calls may be due to a variety of factors e.g. the rise in Respiratory Syncytial Virus infection at this time of year and the use of NHS Direct as an 'out of hours service' over Christmas. It is likely that the 'cold/flu' algorithm is used for respiratory disease other than influenza. Nevertheless, the second peak in the proportion of 'cold/flu' calls during winter 2000/2001 (week 06/01) coincided with both the highest rate in the RCGP consultation rate for influenza and ‘influenza like illness’ and highest weekly number of community based influenza isolates received by the PHLS. The largest age specific increase in the proportion of ‘cold/flu’ calls during the second peak was for 5-14 year olds which also mirrored national surveillance data (RCGP consultations for influenza and ‘influenza like illness’ for 5-14 year olds). This age distribution was different to that of the first peak where various explanations may be responsible for the increase in the proportion of ‘cold/flu’ calls being highest for 15-44 year olds. Where age specific NHS Direct data are available for winter 1999/00, these data also reflect the age distribution of influenza shown by existing national influenza surveillance systems [4].

A potential advantage of using NHS Direct calls for disease surveillance is the ability to aggregate and analyse data on a daily basis. Age specific daily data from 2 NHS Direct sites revealed many un-sustained small peaks in the proportion of ‘cold/flu’ calls. While interpretation of these daily data during a winter characterised by relatively low influenza activity has been difficult, these data may prove to be of use in alerting decision makers to the early stages of a larger epidemic.

Although it has been possible to compare total NHS Direct call rates between winters, it has not been possible to compare ‘cold/flu’ calls for 2000/2001 with 1999/2000 because of the change in clinical decision support systems used during 2000. Consequently, during the winter 2000/2001, the NHS Direct syndromic definition of influenza and ‘influenza like illness’ has not been improved and is based upon a non-specific ‘cold/flu’ algorithm. It has not been possible, using the NHS Direct CAS to extract individual symptoms or clusters of symptoms linked to callers.

Use of NHS Direct data for surveillance of influenza during winter 2001/02 should be improved by two operational changes within the service. Firstly, all 23 NHS Direct sites will be using CAS, enabling the collection of ‘cold/flu’ call data across the whole of England and Wales and enhancing the validity of year on year and regional comparisons.

Secondly, there are plans to centralise national NHS Direct call data within two 'data warehouses' which will help streamline data collection for national work.

Conclusions

NHS Direct is a new and evolving service and analysis of data from future winters is needed to evaluate whether NHS Direct adds value to those routinely available influenza and 'influenza like illness' surveillance systems. Continued efforts will be made to improve the definition of calls for influenza like illness by attempting to identify callers with clusters of symptoms suggestive of influenza.

In future, the effect of Christmas 'out of hours' calls, and the contribution of other seasonal respiratory disease activity to the increase in call rates over the Christmas and New Year period, need to be further elucidated.

The surveillance of age specific NHS Direct 'cold/flu' call data from 6 NHS Direct sites using a single decision support system (covering a population of 16 million), was an important development on the previous year's surveillance and should be pursued.

The ability of NHS Direct data to mirror other surveillance systems, coupled with its timeliness (data available the next day), make it suitable for use in influenza surveillance during the winter of 2001/02. It is hoped that, as NHS Direct reaches more of a 'steady state' in terms of population coverage and uniformity of clinical support systems, it will be possible to begin to construct 'baselines' for the respiratory related disease call data. It should then be possible to provide daily, regional, age-group specific 'cold/flu' data for those involved in managing 'winter pressures'.

References

1. Munro J, Nicholl JP, O'Cathain A, Knowles E**.** Impact of NHS Direct on demand for immediate care: observational study. BMJ 2000; **321**:150-153.
2. 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*, July 2001.
3. Payne F, Jessopp L. NHS Direct: review of activity data for the first year of operation at one site. J Pub He Med 2001; 23(2): 155-158
4. Harcourt S, Smith GE, Hollyoak V, Joseph CA, Chaloner R, Rehman Y et al. Can calls to NHS Direct be used for Syndromic Surveillance? A pilot at three sites using influenza as an example. *Commun Dis Public Health. [IN PRESS]*
5. Dobson R. Health weather forecasts to be piloted in England. *BMJ* 2001;**322**:72
6. Rodman JS, Frost F, Jakubowski W. Using nurse hotlines calls for disease surveillance. *Emerging Infectious Diseases* 1998; 4: April-June
7. Goddard NL, Joseph CA, Zambon M, Nunn M, Fleming DM, Watson JM. Influenza surveillance in England and Wales: October 1999 to May 2000; *Commun Dis Public Health* 2000; 3: 261-6.
8. Fleming DM. Weekly returns service of the Royal College of General Practitioners. *Commun Dis Public Health* 1999; 2: 96-100.

11 Linking syndromic surveillance with virological sampling

Accepted as: Cooper DL, Smith GE, Chinemana F, et al. Linking syndromic surveillance with virological sampling. *Epidemiol Infect 2008: 136(2): 222-4.*

Abstract

Calls to a UK national telephone health help line (NHS Direct) have been used for syndromic surveillance, aiming to provide early warning of rises in community morbidity. We investigated whether self-sampling by NHS Direct callers could provide viable samples for influenza culture. We recruited 294 NHS Direct callers and sent them self-sampling kits. Callers were asked to take a swab from each nostril and post them to the laboratory. Forty two percent of the samples were returned, 16.2% were positive on PCR for influenza (16 influenza A(H3N2), 3 influenza A (H1N1), 4 influenza B) and eight for RSV (5.6%). The mean time between the NHS Direct call and laboratory analysis was 7.4 days. These samples provided amongst the earliest influenza reports of the season, detected multiple influenza strains, and augmented a national syndromic surveillance system. Self-sampling is a feasible method of enhancing community based surveillance programmes for detection of influenza.

Introduction

Effective community surveillance to rapidly identify outbreaks of serious illness is important, particularly given the likelihood of emerging infectious disease and the threat of biological or chemical terrorism. In Europe surveillance has traditionally been performed by general practitioners recording the number of patients presenting with specific symptoms and collecting samples for laboratory testing [1]. In addition, in the UK calls to a national telephone health help line (NHS Direct [2] are used for syndromic surveillance, aiming to provide early warning of non-specific rises in community morbidity [3]. A key limitation of this syndromic approach is the lack of medical or laboratory confirmation of diagnosis.

During the winter of 2003/2004, a pilot study demonstrated the feasibility of community based virological sampling carried out by NHS Direct callers, without the intervention of a health care worker [5]. However, the work was conducted using a small sample (67 callers) and when it was known that influenza was already circulating in the community. We report a further study which investigated whether self-sampling by NHS Direct callers could provide viable samples for influenza surveillance. Also, given the availability of near patient tests for influenza [6], this study also examines the acceptability of self-sampling by NHS Direct callers themselves.

Methods

Between 1st November 2004 and 11th February 2005 callers over the age of 15 years to three NHS Direct sites (Hampshire and Isle of White; West Midlands; South Yorkshire and South Humberside), who sought medical advice about reported 'cold/flu' symptoms and gave their consent to being contacted, were sent a specimen kit in the post. The kit included an information sheet, two nasal swabs, viral transport medium, instructions, appropriate packaging and a short questionnaire. Callers were asked to take a swab from each nostril and return the swabs by post to the national influenza reference laboratory of the Health Protection Agency. The swabs were tested by multiplex PCR for influenza viruses. Material found to be positive by PCR was cultured for viable virus isolation. We used univariate analysis to test the association between both response rate and positivity rate, and a range of call and sampling details.

Results

During the study period the three NHS Direct sites received 244,664 calls for all syndromes, 1,817 calls classified by nurses as ‘colds and flu’, from which 294 callers were recruited for the study. Sixty-one percent of those recruited were female and 50% aged between 25-44 years, reflecting the demographic profile of demand for NHS Direct. Of the 294 sampling kits sent to NHS Direct callers, 142 samples (48%) were returned. The mean time between the NHS Direct call and swabbing was 4.1 days, between the NHS Direct call and laboratory analysis 7.4 days with the range 3-27 days. There were no significant differences in these mean times between positive and negative influenza samples. Callers of 45 years and over were significantly more likely to return swabs to the laboratory ($p<0.001$). Seven out of 141 callers who returned a questionnaire reported minor problems taking the swabs (e.g. ‘spilt the transport diluent’, ‘dropped swabs on floor’).

Table 1: Laboratory results of the 142 NHS Direct samples returned to the national reference laboratory (1st November 2004 to 11th February 2005).

Laboratory result	Number of samples
No virus detected	111
Virus detected	31
Influenza	23
Influenza A (H1N1)	*3*
Influenza A (H3N2)	*16*
Influenza B	*4*
RSV	8
RSV A	*2*
RSV B	*6*
Total samples	**142**

Twenty-three of the NHS Direct specimens (16.2%) were positive on PCR for influenza and eight for RSV (5.6%) (Table 1). Three Wellington/1/2004-like influenza A (H3N2) (the predominant strain in the UK during the winter (4)) and one Shanghai/361/2002-like influenza B virus were recovered following the culture of 4 of the positive specimens. The NHS Direct samples included the 2nd community sample of influenza A (H1N1), 4th of influenza A (H3) and 1st influenza B sample tested by the national reference laboratory during the 2004/2005 influenza season. The overall positivity rate of the NHS Direct samples (16%) was lower than that of the established HPA virological surveillance scheme (26%), although the peak positivity for both schemes was during the first week of February 2005.

Conclusions

This study has demonstrated that people can self sample in a reasonable time frame and that these samples provide good viability for antigenic characterization and molecular detection. This reduces the requirement for the medicalisation of sample taking. This study tested a mechanism where by self-testing by callers could augment syndromic surveillance. Self-sampling by NHS Direct callers provided amongst the earliest reports of influenza circulating in the community and detected multiple strains of the virus.

The majority of the callers who returned samples reported no problems in taking the test. This supports evidence regarding the acceptability of self-sampling in other areas (e.g. sexually transmitted infections [6]). We suggest that this methodology will allow novel approaches to be developed for surveillance of infectious disease using near-patient tests [7], and antimicrobial resistance. The detection of RSV highlights the schemes potential to detect other viruses, and it would be worth exploring possibilities for surveillance of gastrointestinal infections with known surveillance bias (e.g. norovirus [8]), or where vaccines are being considered (e.g. rotavirus [9]).

References

1. Fleming DM, Schellevis FG, Paget WJ. Health monitoring in sentinel practices networks; the contribution of primary care. *Eur J Public Health* 2003; 13(3 Suppl): 80-4.
2. NHS Direct online. http://www.nhsdirect.nhs.uk/ 26th May 2006].
3. Baker M, Smith GE, Cooper D, Verlander NQ, Chinemana F, Cotterill S, Hollyoak V, Griffiths R. Early warning and NHS Direct a role in community surveillance? *Journal Pub Hlth Med* 2003; 25(4): 362-368.
4. HPA National Influenza Season Summary. Influenza / Respiratory Virus Team. HPA Centre for Infections. 2005. http://www.hpa.org.uk/infections/topics_az/influenza/seasonal/activity0405/text/summary.pdf [26th May 2006].
5. Reingold A. If syndromic surveillance is the answer, what is the question? *Biosecur Bioterror* 2003;1(2):77-81.
6. Wiesenfeld HC, Lowry DL, Heine RP, Krohn MA, Bittner H, Kellinger K, Shultz M, Sweet RL. Self-collection of vaginal swabs for the detection of Chlamydia, gonorrhea, and trichomoniasis: opportunity to encourage sexually transmitted disease testing among adolescents. *Sex Transm Dis* 2001 Jun;28(6):321-5.
7. Harnden, A. Brueggemann, S. Shepperd, J. White, A. C Hayward, M. Zambon, D. Crook, and D. Mant. Near patient testing for influenza in children in primary care: comparison with laboratory test**.** *BMJ* 2003; 326: 480.
8. Wheeler GW, Sethi D, Cowden JM, et al. Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. *BMJ* 2000;318:1046-50.
9. The Pediatric ROTavirus European Committee (PROTECT). The paediatric burden of rotavirus disease in Europe. *Epidemiol Infect* 2006;134(5):908-16.

12 General Discussion

Introduction

In recent years there has been a growth in the number of syndromic surveillance systems, many of which operate in the USA. A common aim of these schemes is to collect and analyse pre-diagnostic information in order to provide early warning of rises in disease, and estimate the health status of the community. Since 2001, the UK Health Protection Agency has used data from a nurse-led health helpline (NHS Direct telehealth system) for syndromic surveillance purposes. The purpose of this thesis is to determine what the contribution of NHS Direct data are to health protection surveillance in England and Wales? This has been addressed through a series of papers exploring the utility of call data for providing outbreak detection and early warning of rises in disease. A descriptive study of the socio-demographic profile of NHS Direct callers has been included in order to examine the representativeness of the data for surveillance purposes. Additionally, because syndromic data do not contain definitive diagnoses, only suggestions of potential causative pathogens or environmental triggers, two validity studies are reported. The first tests the feasibility of whether self-sampling by NHS Direct callers can provide diagnostic samples for laboratory testing. The second estimates the contribution of a range of respiratory pathogens to the volume of NHS Direct calls.

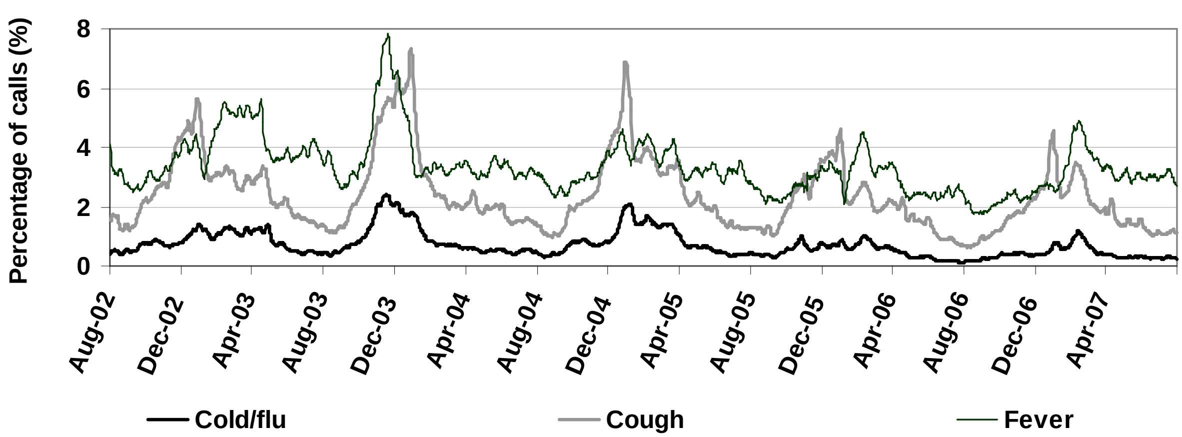
In this final chapter the methodology and operational findings of the NHS Direct syndromic surveillance system (outlined in the proceeding chapters) are examined with respect to the fundamental principles of public health surveillance (HPA, 2005). Also discussed are the key findings of other published work to which the author has contributed to, specifically the impact of environmental health triggers on NHS Direct calls, and a qualitative evaluation of the usefulness of the NHS Direct syndromic surveillance system for its users. Following this there is a discussion of what additional public health value is derived from using NHS Direct call data, and other types of syndromic data, for surveillance purposes. Finally, recommendations are made for the UK Health Protection Agency and NHS Direct, the research community, and those considering using telehealth data for surveillance purposes in other countries.

Main findings

Respiratory infections

There are approximately 300,000 NHS Direct calls per year classified as cough, fever, difficulty breathing or 'cold/flu' syndrome. Annual peaks of difficulty breathing calls (1.1% of total annual calls) and cough calls (2.4%) occur consistently in late December, with peaks in cold/flu (0.7%) and fever calls (3.4%) between November and April. Of these syndromes cough calls exhibit the clearest seasonality (Figure 1), mirroring the sharp late December/early January peaks in RSV and bronchitis in the UK. The seasonality of cold/flu and fever calls is less predictable with the timing of call surges determined by the timing of influenza periods, and to a lesser extent by the RSV season. Our study to estimate the relative contribution of respiratory pathogens to NHS Direct calls concluded that influenza and RSV were responsible for at least half the seasonal variation in NHS Direct respiratory calls. These effects occurred during short lived winter periods, as opposed to an estimated year round contribution of pneumococcol disease and rhinovirus (still with December/January peaks).

Figure 1 - Respiratory syndromes as a percentage of total daily NHS Direct calls in England and Wales (2002-2007).



By examining NHS Direct syndromes individually this work has determined the most likely pathogenic causes of rises in individual symptoms. In turn, this clarifies and supports interpretation within syndromic surveillance reports sent to public health teams in the UK. For example, NHS Direct 'cough' appears to be a generic marker of RSV, bronchitis and total respiratory infections during a December peak. Minor spring peaks in cough calls implicate parainfluenza. Alternatively, NHS Direct cold/flu calls (all ages) and fever call (children only) are a timely proxy indicator for seasonal increase in influenza viruses. Although other respiratory pathogens and other non-infectious causes contribute to NHS Direct respiratory calls the contribution of these causes cannot be estimated due to a lack of laboratory data (e.g. coronavirus, human metapneumovirus) from which to model an effect, or due to lack of significant correlation within our regression models (e.g. legionella). The importance of influenza, RSV and *S. pneumoniae* in determining call rates supports other work that has estimated the impact of these diseases on GP consultation (Melegaro et al., 2006) and hospital admissions (Muller-Pebody et al., 2002; HPA, 2005). For example, s *pneumoniae* is estimated to cause 30% of NHS Direct cold/flu calls, 25% of community acquired pneumonia (CAP) and otitis media seen by GPs, and 29% of hospital admissions with CAP (Hawker et al., 2001).

The seasonal fluctuation in cold/flu and fever calls closely mirrors GP consultations for influenza-like-illness. NHS Direct and GP incidence data were modelled so as to derive numerical thresholds values for influenza surveillance. These thresholds (cold/flu: 1.2% of total calls, fever (5-14yrs): 9% of total calls) were exceeded two weeks in advance of existing clinical thresholds based on GP incidence (Goddard et al., 2003), providing accurate early warning of seasonal influenza activity. This advance warning was demonstrated during 4 of the 5 winters studied. By using age-group specific NHS Direct data it was possible to distinguish which age-group specific syndromes were useful for different influenza sub-types; e.g fever (5-14yrs) for the influenza B 2005/2006 season, cold/flu calls (all ages) for the relatively early influenza A(H3N2) 2003/2004 season.

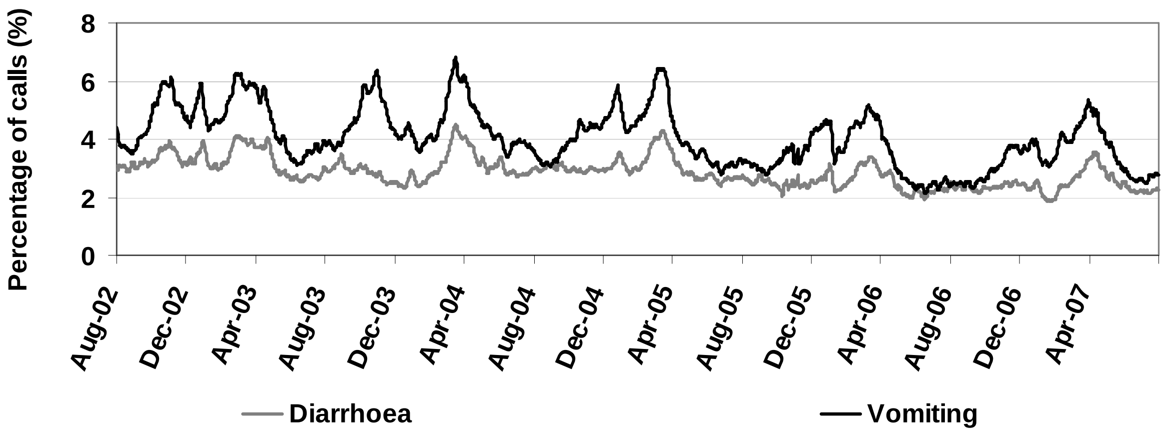
A method of supporting syndromic thresholds, and community surveillance in general, was demonstrated by the pilot study to investigate the feasibility of NHS Direct callers self-sampling for influenza and RSV laboratory testing. After a cold/flu call it took on average seven days for a caller to be sent a testing kit, self-sample, mail the swabs to the laboratory, and for a laboratory test to be made available. Of the 142 samples taken as part of the pilot, 23 (16%) tested positive on PCR for influenza (4 also by culture), similar to the regression model estimate of 22%. These specimens demonstrated that self-sampling was a feasible method of community influenza sampling (despite samples spending a week in transit), with potential value for surveillance of other diseases (e.g. norovirus). The specimens provided some of the earliest influenza reports of the season and augmented the purely syndromic approach.

Analyses of respiratory calls may serve as a useful proxy for influenza surveillance, not only on a national but also on a local level. Spatio-temporal analyses of fever calls for the 5-14yr age group provided a timely and unique description of the evolution of a national influenza outbreak (Chapter 8). Two distinct waves of significantly raised levels of NHS Direct fever calls were identified during a retrospective study of weekly call data. The first wave originated in the North West of England during late November 2005 and spread south eastwards; the second wave began in Central England during mid-January 2006 and moved southwards. The timing and geographical location of these waves of fever calls were similar to the national influenza B outbreak occurring in the same winter, which at the time was not mapped or described to the same level of geographical detail. Telehealth data are thus a potentially useful tool for influenza surveillance at both national and local level. Such a tool has utility for local services in alerting heath care providers of increased seasonal influenza activity, and potentially in the event of an influenza pandemic.

Gastrointestinal infections

The incidence of both diarrhoea and vomiting calls exhibits a double peak during winter and spring (Figure 2). The age and seasonality distribution, however, differs between syndromes. Diarrhoea calls (2.9% of total calls) peak during late December (adults only), followed by a longer higher peak during March and April (almost exclusively confined to children under 5). Vomiting calls (4.1% of total calls) exhibit the same double peak, although the seasonal variation is greater and is dominated by calls about children under 5 years.

Figure 2 - Gastrointestinal syndromes as a percentage of total daily NHS Direct calls in England and Wales (2002-2007).



The seasonality of NHS Direct diarrhoea and vomiting calls is largely driven by common viral rather than bacterial infectious intestinal disease. This statement is supported by statistical evidence derived from multiple linear regression models to estimate the contribution of enteric pathogens to NHS Direct gastrointestinal calls for the 0-4 year age-group (Cooper, 2007; Harris et al., 2007). These models estimated that rotavirus accounted for 27% of diarrhoea and vomiting calls, almost exclusively between February-April. A similar contribution of rotavirus has been estimated for GP consultations (25%) and A&E admissions (20%) (Harris et al., 2007). Norovirus is estimated to cause 15% of vomiting calls for the 0-4 year age group, with a peak contribution between January and April during periods of raised outbreak reporting. By comparing time-series of vomiting calls and norovirus laboratory reports, working NHS Direct vomiting threshold values have been established for norovirus surveillance (Loveridge et al., 2007). National proportions of vomiting calls of 4% and 3% for all ages and ≥5yrs respectively have been used prospectively during winter 2007 to alert health protection teams to the beginning of periods of increased norovirus activity nationally. Age-group specific syndromic data (about hospital visits) has also proven useful in understanding the seasonality of gastroenteritis in hospital settings, also demonstrating winter and spring peaks (Olson et al., 2007).

In the UK, bacterial pathogens commonly peak in summer (campylobacter in July, salmonella in August, and *E Coli* O157 in September (Hawker et al., 2001)). These periods of raised incidence have no significant impact on the seasonality of NHS Direct gastrointestinal calls. This is because although common bacterial enteric pathogens are detected in high numbers by laboratories, community incidence is far lower than for the common GI viruses' rotavirus and norovirus. A national study of infectious intestinal disease in the UK estimated that for every laboratory isolate reported to national surveillance, there were an estimated 3.2 cases of salmonella in the community, 7.6 cases of campylobacter, 35.1 cases of rotavirus and 1,562 cases of norovirus (Wheeler et al., 1998). NHS Direct data therefore represent an opportunity to monitor the incidence of community infections, in this case rotavirus and norovirus, not accurately described by other surveillance systems. This also presents an opportunity to identify GI viral outbreaks that cause widespread morbidity in the UK. The difficulties associated with detecting GI outbreaks on a local level are discussed later.

GP consultations about infectious intestinal disease (IID) exhibit a regular peak in children during March which is again suggestive of a viral rather than bacterial cause. In addition to this spring peak, NHS Direct diarrhoea and vomiting calls also peak during December. This pattern was not explained by our regression modelling or the seasonality of common enteric pathogens recorded by laboratory based surveillance systems. Astrovirus causes mild diarrhoeal illness in children under 5 and exhibits a peak in laboratory reports during December and January, but has a low community incidence in industrialised countries (Bon et al., 1999; De Wit et al., 2001). Astrovirus was not a significant predictor variable in the regression models. Further sampling by NHS Direct callers for GI pathogen detection, and examination of other factors (e.g. lifestyle) is required to explain this December peak.

Environmental health effects

The temporal variation in NHS Direct calls is sensitive to environmental determinants such as heat, pollen, and possibly ozone and thunderstorms. Leonardi et al have shown that NHS Direct 'heat and sun stroke' calls, and to a lesser extent fever calls (0-4yrs), rise during heat waves (Leonardi et al., 2006). During summer, the syndromic surveillance team have monitored the use of NHS Direct heat and sun stroke algorithms as part of the UK Department of Health heatwave plan (Department of Health, 2004). The number of daily 'heat/sun stroke' calls was too small to accurately quantify the size of the heat effect (daily 2006 heat wave peak = 115 heat stroke calls in England and Wales), but real-time analysis of call data provided a timely marker of the existence of heat related illness occurring in the community. For example, the impact of heat on call rates about a wider range of syndromes is currently being studied to improve measurement of the heat wave effect. Real-time automation of geographical mapping of heat related calls could more accurately pinpoint area to target rapid interventions, as has been demonstrated by surveillance during longer running initiatives elsewhere (Weisskopf et al., 2002).

During August 2002 there was a statistically significant rise in difficulty breathing calls at NHS Direct call centres in the Thames basin and East Anglia regions of England. It was not possible to determine the cause of the individual calls, although during the days prior to this rise high ozone levels and thunderstorms (often associated with raised hospital admissions for asthma) occurred in this part of England. In this case syndromic surveillance data provided support of anecdotal rises in asthma admission in hospitals. Finally, during successive June periods national and regional peaks in NHS Direct 'eye problems' calls (5-44 year age-group) have occurred during peaks in grass pollen counts (Cooper, 2007). The exact timing and magnitude of these peaks vary according to the weather and timing of grass pollen seasons and have been communicated to the health service via email alerts.

Operational findings

The rationale of using NHS Direct data for surveillance is that rapid collection and analyses of these data will lead to timely identification of trends in community morbidity in England and Wales. Initially data from a few pilot NHS Direct sites were used for influenza-like-illness surveillance. When national NHS Direct data became available, soon after the events of 9/11, the aims of the surveillance system were expanded to include identification of an increase in syndromes indicative of infections caused by the deliberate release of a biological agent. A multidisciplinary group of doctors, epidemiologists and scientists defined a set of prodromal and syndromic conditions associated with potential bio-terrorist agents (Baker et al., 2003). In the absence of biological attacks on the UK it has not been possible to evaluate whether this aim has been met. This lack of opportunity (thankfully) has led to further development of the system's aims: to identify increases in syndromes indicative of naturally occurring infections (other than ILI) and environmental influences on health. To this end, the system has routinely monitored daily data about 10 syndromes reported to 23 NHS Direct call centres between 2001 and 2006, and for 10 Strategic Health Authorities from 2006 onwards.

During the early operation of the surveillance system (December 2001-February 2003), a total of 1,811 statistical exceedances ('signals') occurred, of which 126 (7%) were investigated and 16 (1%) resulted in public health 'alerts' being issued. The large amount of signals meant that the surveillance team only had the resources to investigate those that were considered unusual in some way (e.g. affecting a particular age-group, or involving a high proportion of calls with serious outcomes such as a ambulance dispatch or emergency department referral). The main reason that the 126 investigations did not progress to alerts were:

1) the observed increase in calls was a single-day alarm only and dropped the following day (46% of investigations),
2) duplicate call records caused the alarm (20%), and
3) the call data did not cluster geographically (15%).

The problem of duplicate call records was removed from the system, but there persists a high number of statistical signals, few of which progress to investigation and alerting. This high signal to noise ratio does not cause undue pressure on the resources of the surveillance team because the signals are automatically generated by computers and are often dismissed after only a brief initial investigation.

Published results of 12 months operational NHS Direct alerts (Baker et al., 2003) reported that 3 out of 7 alerts had positive findings, detecting local rises in rotavirus and norovirus, and a national rise in influenza. It was not possible to take samples from these callers as there was (and still is) no *routine* mechanism for doing so. The surveillance 'alerts' were communicated via email or phone call within 1-2 days of the surge in calls being identified. Despite refinement of the methods used to derive statistical exceedances (the development of control charts), the proportion of signals leading to positive alerts of local health protection events has not improved. This has caused those charged with analysing and interpreting data to concentrate on the areas where successes *have* been achieved and telehealth data have provided demonstrable added value; namely the surveillance of influenza, other viral respiratory and gastrointestinal disease, and heat related illness on a regional and national level.

The flexibility of the surveillance system is such that that data have been extracted and analysed three times a day during periods of increased perceived risk (e.g. after the London ricin arrests (January 2003) and London Bombs (July 2005)). More recently it has been possible to extract daily data linked to postcode districts. This means that the surveillance system provides the only UK source of 'real-time' intelligence, from which the surveillance team has developed a responsive incident surveillance capability. This capability was used during July 2007 for surveillance of the health effects of major floods affecting large parts of England. NHS Direct data were extracted from the national database for two baseline periods (just prior to the floods, and for the same period during previous years) and compared to prospective data on a daily basis. A 40% rise in total NHS Direct calls in the Gloucester flood region was observed, caused in part by a rise in people seeking health information and advice. There were no significant short term rises in gastrointestinal, ear or respiratory complaints, asthma, psychological problems or injuries. These analyses were used to brief a national incident team and the Government Cabinet Office on the health effects of the floods.

Ascertaining individual cases of public health importance

A persistent problem for syndromic surveillance systems is the inability to distinguish the syndromic cases that cause the statistical excess (and may originate from a specific outbreak) and the cases that comprise the baseline level (Murray et al., 2007). Although the NHS Direct syndromic surveillance system routinely monitors disease on a population level, the following examples demonstrate how, on occasions, this work has identified potentially significant individual cases of disease. Firstly, the pilot study to develop a mechanism for self-sampling by NHS Direct callers themselves (Chapter 11) successfully collected samples positive for influenza A, B and C and RSV type I and II. Although self-sampling by callers is still not routine, it offers the possibility of an alternative source of specimens in exceptional circumstances (e.g. pandemic flu, or an unexplained rise in illness), or for routine community based surveillance. The need for diagnostic linkage of data has been recognised elsewhere with evidence that by filtering syndromic data to only cases with linked laboratory tests, sensitivity for outbreak detection may be improved (Babin et al., 2007).

A surveillance ‘flag’ on the NHS Direct call handling software automatically highlights single callers triaged using the ‘lumps’ algorithm and reporting painless black lesions (i.e. suggestive of cutaneous anthrax syndrome). During January 2005 this flag identified 3 callers each of which was a false positive anthrax case (1 post-BCG vaccination lump, 1 post hospitalisation venepuncture lump, 1 possible ganglion). Although false positives were identified by the flag these cases demonstrated that it was possible to explore additional sensitivity for bio-surveillance, over and above the NHS Direct ‘lumps’ syndrome monitored routinely.

Finally, in exceptional circumstances, surveillance team on-call doctors have contacted NHS Direct callers to discuss their illness. For example in January 2002 there was an unusually sharp and statistically significant rise in callers reporting fever calls in children. The on-call doctor phoned a sample of callers and concluded that the illness was self limiting and respiratory in nature. The rise in fever calls was later found to be associated with the emergence of the antigenic shift variant influenza A(H1N2) that was not initially identified by laboratories or clinical ILI surveillance systems (Goddard et al., 2004). Caller contact is rare, however, and an ecological approach is taken with the majority of surveillance signals investigated.

Discussion

Representativeness of the system

To effectively interpret NHS Direct surveillance data, it is necessary to understand how representative callers are of the UK population, health care usage in general, and the UK disease burden.

The UK population and health care usage

One quarter of the population use NHS Direct (Knowles et al., 2006) suggesting that there may be a potential population bias in the caller sample. This is borne out by total usage of NHS Direct (8 million calls per annum) which although comparable to GP out-of-hours calls (8 million (Salisbury, 2000)) and hospital accident and emergency department visits (14 million), is a fraction of the 190 million GP consultations (Rowlands and Moser, 2002). If pharmacy sales are also added to this list, NHS Direct is the fifth most used health service in the UK. US research supports this position placing telephone advice low in the list of health seeking behaviour (Metzger et al., 2004) and equal in usage to internet advice services (Hakre et al., 2007). As opposed to GP services and A&E Departments, which have been used by the UK population for decades, NHS Direct has only operated since 1998. This may result in differential uptake between socio-demographic groups and geographical areas, now discussed.

NHS Direct currently receive a disproportionately high number of calls about children under 4yrs (four times the total call rate) and women of child bearing age (twice the male call rate). Both these groups traditionally have a high health service utilisation in the UK (Rowlands and Moser, 2002). Old people, who have greater GP and hospitalisation rates, use NHS Direct less than their population composition would suggest. This age and gender bias has been constant since NHS Direct's inception (Munro et al., 1998) and is consistent with large public telephone triage systems in New Zealand (St George et al., 2006), Australia (Turner et al., 2002) and Canada (Rolland et al., 2006).

There are also significant differences in NHS Direct call rates between localities, with a seven-fold variation in total call rates between Primary Care Trusts (PCTs) in England, and four-fold variation in rates between administrative wards within a single PCT. These disparities are only partially explained by social deprivation: call rates are 40% higher in areas where deprivation is at or just above the national average (Chapter 3; Burt et al., 2003). Salisbury's work calculated that GP out-of-hours services generate on average 160 calls per 1,000 per annum (Salisbury, 2000). The introduction of NHS Direct has been associated with a small reduction in calls to GP cooperatives (Munro et al., 2005). It is, therefore, reasonable to assume that there is a trade off between usage of the two services, and well established local GP out of hours services may restrain demand for NHS Direct. In conclusion, it appears that local arrangements for provision of GP out-of-hours services, social deprivation, and also the length of time that the local NHS Direct site has been in operation (Knowles et al., 2006) are major determinants of local call rates. Although further research is required to fully quantify the impact of these and other potential factors, the NHS Direct surveillance system is broadly representative of primary care usage of those aged below 65 years.

<u>The UK disease burden</u>

The estimates of the contribution of specific infectious disease pathogens (e.g. rotavirus, *s pneumoniae*) to NHS Direct syndromic calls are similar to GP and hospital estimates, both for adults and children. Additionally, in some instances NHS Direct call data may provide a more accurate description of rises in childhood disease incidence than other surveillance sources. For example distinct rises in vomiting calls (0-4yrs) and fever calls (5-14yrs) have preceded rises in norovirus outbreaks mainly reported in institutional settings, and influenza-like-illness reported to GPs respectively. It is also possible, if necessary, to monitor illness during weekends and bank holidays when GP surgeries are closed.

Old people, who suffer higher rates of chronic conditions and have acquired immunity to certain infections, are less likely to phone NHS Direct. Call rates from old people are relatively low and there is less seasonal variation in the incidence of calls suggestive of infections. Consequently, the syndromic surveillance system appears unrepresentative of the disease burden in this age group, and provides little added value over sentinel GP surveillance systems. Elderly people, however, are the UKs largest vulnerable group. Influenza A(H3N2) viruses, for example, are known to cause high morbidity in the elderly and are associated with concurrent high bronchitis and hospital admission rates (Goddard et al., 2000). This demonstrates the need for complementary surveillance systems, analysing and presenting both GP and NHS Direct data. The Primary Care Surveillance Unit of the Health Protection Agency has such a remit, and is thus able to monitor influenza strains (and other acute public health problems) in all age groups.

In the UK, it is well established that deprived groups have higher rates of hospital admissions and GP consultations, both for all conditions (Carlisle et al., 98; Saxena et al., 99) and for common infections (Fleming and Charlton, 1998; Olowokure et al., 1999; Hawker et al., 2003). This is consistent with higher respiratory, gastrointestinal and total NHS Direct call rates in more deprived areas (Cooper et al., 2002). The utility of the system is greatest for acute common infections with high childhood incidence rates (e.g. RSV, rotavirus and norovirus). Less common infections (e.g. HIV or tuberculosis) often exhibit a high incidence in socially excluded groups or distinct ethnic populations. Although the infected individuals may phone NHS Direct, surveillance of rarer chronic infections is perhaps more suited to schemes where a definitive diagnoses is possible, and where daily surveillance output is not an important measure of system performance.

The studies presented in this thesis have employed two methods to account for geographical and age-specific variation in call rates. Firstly, and in common with other syndromic surveillance practice (Heffernan et al., 2003), age-group specific case counts have been used as a denominator rather than the traditional population counts. This means that syndromic indicators are proportions rather than rates, enabling a valid comparison of call incidence between geographical areas with differing socio-demographic characteristics. It also controls for short term fluctuation in population call rates caused by day of the week and bank holiday effects. Secondly, spatio-temporal analysis of fever and vomiting calls, used to examine disease diffusion (chapter 8), established expected weekly counts for each area (in this case Primary Care Trusts). It was then possible to detect local variation in fever and vomiting call incidence. This approach has also been used successfully for surveillance of specific infections (e.g. Shigella (Jones et al., 2006), West Nile Virus (Mostashari et al., 2003)) and syndromic emergency department data (Heffernan et al., 2003). Population change, which could have caused a local rise in calls and misleading clusters, was considered negligible as the study period covered a single year. For longer term monitoring a moving historical baseline should be employed to account for population change.

Sensitivity, specificity and positive predictive value

Sensitivity measures the ability of a test to detect an individual with disease. This definition can be interpreted in two different ways with regards to syndromic surveillance. Firstly, sensitivity could refer to the ability of an individual NHS Direct call to signify an individual disease (e.g. influenza infection). Secondly, sensitivity could refer to the ability of a surveillance signal (significantly high number of syndromic calls) to signify a real outbreak. Regarding the first definition, this work has used two methods to define sensitivity. The first, self sampling by NHS Direct callers, showed that 16.2% of adult ‘cold/flu’ callers tested positive for influenza out of a sample of 294 ‘cold/flu’ calls during winter. This pilot had a lower sensitivity than virological sampling initiated by GPs during the same season (38%) (Cooke et al., 2004). The second (Chapter 7) used a statistical model to estimate the sensitivity of ‘cold/flu’, cough, fever, diarrhoea and vomiting calls to various infectious disease pathogens. The model results (discussed already) provide clues to the aetiology of rises in calls throughout the year aiding interpretation of data.

The second interpretation of sensitivity concerns outbreak detection. This is more pertinent to syndromic surveillance systems, as they routinely monitor disease on a population rather than individual level. To test the sensitivity of syndromic surveillance signals one can employ a gold standard for comparison (e.g. outbreak reporting schemes or clinical surveillance schemes), or conduct a simulation study of outbreak conditions. The sensitivity of our control chart methodology to detect local outbreaks of disease is low. This was demonstrated by the simulation study (Chapter 5) which concluded that the current surveillance system would have a 4% chance of providing early warning of a local cryptosporidiosis outbreak if 5% of those infected reported their illness to NHS Direct (specificity 99.5% - false alarm rate 0.5%). Other simulation studies provide detection rates for hypothetical bio-terrorist events such as anthrax release. Nordin found that if 4% of anthrax cases (from a simulated mall outbreak) reported to a physician, there was a 5% probability of detecting the anthrax attack (specificity 99% - false alarm rate 1%) (Nordin et al. 2005). Not surprisingly, and similar to the NHS Direct simulation, the probability of early detection increased rapidly once reporting rates increased (e.g. for a 40% reporting rate there was a 75% chance of detection of the anthrax attack within 7 days of the release, and 10% chance of detecting the cryptosporidiosis outbreak before traditional means). An unrealistically high reporting rate will generate a false impression of the power of syndromic surveillance, however, so it is necessary to model a range of scenarios.

Another important consideration is the detection filter used for generating surveillance signals. The NHS Direct cryptosporidiosis outbreak simulation used single day and two-day linear temporal filters (i.e. the excess during each of the two previous days is of equal importance) and the anthrax study 1, 2 and 3 day linear filters. Reis has demonstrated that non-linear 7 day filters, where each preceding day's data is of diminishing importance, can improve sensitivity under simulated conditions (Reis et al., 2003).

Simulation studies inform epidemiologists of the expected sensitivity of a system. Prospective results provide evidence of the usefulness of a system and how human judgement (i.e. which signals to follow? which to ignore?) also influence outcome. In a well argued critique of syndromic systems, Reingold suggests that even when statistical signals are detected, they are seldom investigated, diagnostic testing is not necessarily initiated, and there is an associated cost of investigating false alarms (Reingold, 2003). The aims of syndromic surveillance, however, have evolved since this article was written. Those responsible for investigating signals are now more realistic about the public health significance of data aberrations. Most single day signals are dismissed with little additional investigation and associated cost. Only when a signal progresses to an alert can the utility of the system be demonstrated. Public health teams can then learn what type of local and national events syndromic surveillance can and cannot detect. This daily 'system practice' also helps to maintain the systems infrastructure, the public health links and communications, and ensure readiness during real emergencies.

The decision whether to further investigate and ultimately report a significant rise in calls (signal) is not purely statistical. The personal experience of the epidemiologist, other surveillance and media intelligence, the current political climate, and staffing levels, will also influence this decision. The positive predictive value of investigations is perhaps more relevant to syndromic surveillance than the sensitivity of signals per se. Limited previous success in detecting outbreaks may also slow down initiation of an investigation (Murray et al., 2007). Various attempts have been made to help epidemiologists prioritise which statistical anomalies to pursue further. A framework for prioritisation based mainly on statistical attributes of the data is being explored by the national US surveillance systems (BioSense), to help identify potentially important anomalies (Martin et al., 2007). A more qualitative approach has examined the epidemiologist's decision making process and concluded that anecdotal evidence from health care providers, corroborative data from other systems, the magnitude and duration of the increase, and whether the increase is expected or unexpected, all influence the decision making process (Hurt-Mullen and Coberley, 2005).

Added value of the NHS Direct syndromic surveillance system

Early warning and outbreak detection.

The early warning potential of telehealth data may be aided by three factors. Firstly, the public may contact NHS Direct before visiting GPs or hospitals to gain health advice. Secondly, NHS Direct may be the only point of contact some people have with the health services due to the self-limiting nature of their illness (i.e. morbidity 'under the radar' of other systems). Thirdly, daily NHS Direct reporting is more timely than reporting from GP surveillance systems where the Monday to Sunday period is routinely reported the following Wednesday. Laboratory data from community virological schemes may take up to 10 days from specimen collection to reporting. Despite these potential advantages, this work has not demonstrated surveillance utility for local outbreak detection (e.g. food borne outbreak of salmonella, school outbreaks of influenza-like-illness, cryptosporidiosis outbreaks caused by a contaminated water supplies).

The poor sensitivity and low signal to noise ratio on a local level may be due to a combination of the following factors:

1. Total NHS Direct call rates are one thirtieth of GP consultation rates, meaning only a fraction of ill people phone NHS Direct.
2. The unspecific nature of syndromic cases mean outbreaks cannot be identified above a background syndromic baseline.
3. The relatively low geographical granularity of the surveillance (routine data analysis is for spatial units of over 1 million people).
4. An inadequate statistical and epidemiological methodology.

Since the surveillance system became operational the NHS Direct site (average population: 2.5 million) and the Strategic Health authority (SHA) (5 million) have been the spatial unit of analyses. Consequently the surveillance system is not expected to detect very local outbreaks of disease. However, even if the spatial unit was small it remains unlikely that the system would be any more successful at identifying local anomalies. By way of example, NHS Direct receive approximately 8 million calls per year in England and Wales. This equates to an average total of 4 calls per day in a town of 10,000, and one call every 5 days about vomiting. Therefore, however finely tuned the statistical algorithms are, a rise in illness in a town of 10,000, i.e. an outbreak, is unlikely to be detected.

As the population of the spatial unit increases, the number of calls per day rises to a level where it is possible to attain statistical significance with relatively small increases above baseline levels. Despite unspecific syndromic case definitions and relatively high background levels of syndromic calls, the control chart methodology has provided early warning of community wide health protection issues. This demonstrates that significant numbers of people with infectious or environmentally triggered disease contact NHS Direct. Spatio-temporal analyses of fever calls to a Primary Care Trust (PCT) level (average population 250,000) demonstrated that it was possible to detect a sub-regional rise in fever calls, consistent with evidence of the origins of a national influenza B outbreak. In summary, significant rises in calls at a national and sub-regional level identify important public health events but call volumes at a sub-PCT level appear too small to be useful in detecting local outbreaks.

Providing reassurance

Reassurance that a rise in illness is not occurring (during times of perceived higher risk) has been suggested as one of the main benefits of syndromic surveillance (Sosin and De Thomasis, 2004). This reassurance is only valid, however, if it has been demonstrated that surveillance can prospectively detect a real increase in illness. The NHS Direct syndromic surveillance system is sufficiently sensitive to detect regional and national events so may also provide some reassurance that a widespread rise in disease has *not* occurred. Below this level reassurance is harder to quantify. For example, during

December 2005 there was a major oil depot explosion in Hemel Hempstead, South East England. Enhanced surveillance of NHS Direct calls was conducted. Three weeks after the explosion a significant rise in respiratory illness was detected at the call centre nearest the fire (Cooper, 2007). Mapping of these data demonstrated no clustering of calls in the exposure zone down wind of the blast, leading to the conclusion that the likely cause was a seasonal rise in respiratory disease. The small numbers of daily calls involved in this analysis (n=16), however, raises some concerns about the sensitivity of such a local investigation. All that can be concluded is that a large scale rise in illness in the exposure zone was unlikely to have occurred, but that smaller population health effects were not demonstrable (although environmental sampling results found no serious ground contamination). Elsewhere reassurance has been provided by the New York syndromic surveillance system. During March 2003 it was concluded that a rise in febrile illness was not caused by SARS, although importantly, resources were available to investigate each individual case (Steiner-Sichel et al, 2004). The NHS Direct surveillance system, in its current form, would not detect a bio-terrorist event affecting small numbers of people such as the ‘anthrax letters’ of 2001 that killed five and infected 17 others across several US states.

Validating and supporting other surveillance schemes

In addition to the NHS Direct surveillance system there are three main sentinel surveillance systems based on family doctor diagnoses currently operating in England and Wales; the Royal College of General Practitioners Weekly Returns Service established during the 1970s and covering 700,000 people (Fleming, 1999), the Q-Research surveillance systems established in 2003 covering 3 million people, and Q-Flu also established in 2003 and covering 22 million people (Hippisley-Cox et al., 2006). It is possible, therefore, to use pre-primary care (NHS Direct) and primary care data (doctor diagnoses) to track illnesses which may not present to hospitals or for which laboratory specimens are not routinely taken (e.g. influenza). Where as the GP based systems monitor weekly diagnoses and prescriptions, the NHS Direct syndromic surveillance system provides a daily, if less specific, picture of morbidity. A qualitative user evaluation of the usefulness of primary care surveillance systems (discussed later) identified the value of the combined primary care surveillance messages as greater than the sum of the parts (Browne et al., 2007). Those surveyed also highlighted the usefulness of NHS Direct and GP systems for interpreting the significance of local anecdote and media reports. With respect to clinical practice, NHS Direct influenza thresholds provide, along with other systems, clear and consistent messages to the NHS, media and public, about current influenza activity (placed within the context of previous winters).

Usefulness as perceived by users of the system

An evaluation of the utility of syndromic surveillance systems operational in 2002 highlighted the lack of published information about how they facilitated decision making (Bravata et al, 2004). It concluded that there was little scientific evidence to guide those running such systems. In 2004, the US Centres for Disease Control (CDC) published the ‘Framework for Evaluating Public Health Surveillance Systems for Early Detection of Outbreaks’ (Buehler et al., 2004). An important element of the CDC framework was the need to describe system usefulness; defined as the added value and impact of surveillance.

Two formal evaluations of the NHS Direct syndromic surveillance system, involving the author, have been undertaken. The first, a preliminary evaluation using the CDC framework (Doroshenko et al., 2004), was conducted during 2004 and included quantitative and qualitative components. It concluded that the system was representative of the population, conducted real-time surveillance and was capable of detecting ‘high-risk, large scale’ events, at a direct cost of approximately £150,000 per annum. Improvements to flexibility and portability were deemed necessary. The work, however, focussed only partially on the systems usefulness and was undertaken just three years into the project. Since this preliminary evaluation, further quantitative studies have evaluated the systems sensitivity for outbreak and event detection (chapters 5-8). Additional functionality (local call data) and syndromes (heat stroke) have been added to the system.

In 2006, the second formal evaluation of the system was conducted (Browne et al., 2007). This time the evaluation team took a qualitative approach to assess the usefulness of syndromic surveillance to the HPA, the National Health Service, and the Department of Health. The team documented the public health action resulting from surveillance outputs, and explored how these outputs could have been of greater public health use. The users of the system (those that received surveillance alerts and bulletins) were asked to complete a self-administered questionnaire (81 responded) or an interviewer administered semi-structured questionnaire (20 interviewed). In response stakeholders gave practical examples of the usefulness of surveillance bulletins, mainly in the area of acute response. Cited examples of action taken as a result of surveillance alerts were; reporting to incident teams and Primary Care Trusts, input to media messages and question and answer briefings for GPs and the public, supporting national plans (e.g. heat-wave plan), and preparing laboratory staff.

Interestingly, those surveyed suggested that a ‘recommended public health action’ section should be included in surveillance bulletins, alongside the existing interpretation of syndromic trends. The provision of suggested action raises the following issues. It is the role of the surveillance team to advise as well as inform system users; it is acceptable to users to receive this advice; and surveillance teams must be truly multi-disciplinary. In addition, if specific public health actions are recommended then there is a risk of directly attributable inappropriate or costly actions. Consequently epidemiologists must carefully consider the impact of their recommendations.

Syndromic surveillance - an international perspective

Outside the UK, there are reported examples where syndromic surveillance systems have provided value for early outbreak detection. For example, the New York City syndromic surveillance system has provided the earliest indication of city-wide influenza activity and identified city wide increases in gastrointestinal disease in advance of a rise in institutional outbreaks of norovirus (Heffernan et al., 2004; Steiner-Sichel et al., 2004). Also in New York City, syndromic surveillance detected an increase in diarrhoeal illness, prompting an investigation to identify otherwise unreported disease caused by the consumption of bad meat that had spoiled after a power cut (Marx et al., 2005). In Kingston, Canada, surveillance of emergency department GI syndromes led to the early identification of a widespread salmonella outbreak and subsequently helped monitor the effectiveness of public health interventions (Moore et al., 2007). Retrospective spatial analyses of hospital data from the Netherlands identified a space-time cluster of pneumonia matching the time and place of a known legionnaires disease outbreak at a flower show (Wijngaard et al., 2007). The cluster was one of many, however, so it was not possible to tell whether it would, at that time, have led to an investigation. In summary, despite these successes, syndromic surveillance has not consistently demonstrated usefulness for detecting acute *localised* outbreaks. Published evidence supports this position highlighting a lack of utility for, GI outbreak detection (Balter et al., 2005), water borne disease detection (Berger et al., 2006), and outbreaks detected in primary care (Yih et al., 2007).

The first published report of the use of telehealth data for surveillance was a retrospective study demonstrating a 17-fold increase in nurse hotline calls about diarrhoea during a citywide outbreak of cryptosporidiosis in Milwaukee (Rodman et al., 1998). Using telephone triage data from the Washington DC area Henry demonstrated that individual calls can be used to predict respiratory and gastrointestinal final diagnoses (Henry et al., 2004). On an ecological level, Wilson showed that rises in respiratory and gastrointestinal calls to the Scottish national telephone triage system (NHS24) heralded later rises in influenza-like-illness and norovirus reports respectively (Wilson et al., 2006). In the US, studies comparing telephone triage data against CDC influenza surveillance data have conflicting results, one study showing a peak in primary care telehealth data after laboratory and clinical data (Espino et al., 2003), the other showing concurrent peaks (Yih et al., 2006). Both these studies concentrated on examining peaks in data or cross-correlation of entire time-series, however, rather than the early warning potential for heralding important public health events.

In Canada, a major evaluation is underway to explore the usefulness of real-time telephone triage data from Ontario's Telehealth helpline for the State's syndromic surveillance strategy (Rolland et al., 2006). Initial results are promising and support the results of this thesis in a number of ways. When studied retrospectively, increases in telehealth calls suggestive of ILI preceded those of ILI recorded by primary care clinics (Rolland E, 2007) and hospital emergency departments (van Dijk et al., 2007). Importantly, of these three data sources, telehealth data can be collected and analysed in the timeliest way, improving the lead time and early warning potential further. Also, in accordance with NHS Direct estimates, it was influenza and RSV that best explained the temporal variation in total respiratory calls in Ontario (van Dijk et al., 2007). To date there are no published analyses of gastrointestinal calls, or spatial analyses to a local level from Canada. It remains to be seen whether this work will prospectively provide early warning and useful state wide reporting of significant public health events.

The studies reported in this thesis have concentrated on the potential of syndromic data for providing *early warning* of rises in disease. In accordance with experience from other countries, success has been achieved in identifying rises in community wide viral gastrointestinal and respiratory disease. There are additional benefits, however, that may ultimately improve the sustainability and acceptance of syndromic surveillance systems within mainstream public health. Firstly, syndromic data provide a unique source of intelligence *throughout* known outbreaks (Murray et al., 2007, Guerrero et al., 2007). NHS Direct surveillance data have been used to provide regular and timely surveillance updates during national outbreaks and major environmental incidents. Buehler has suggested that syndromic surveillance systems should support public health policy by attempting to measure as well as detect problems (Buehler, 2004). Public health policy goals often require surveillance at their core. Estimates of the number of NHS Direct calls caused by specific pathogens have contributed to a national burden of disease report (HPA, 2005) and have been used to evaluate the potential impact of a rotavirus vaccination strategy (Harris et al., 2007). Finally, a commonly cited benefit of syndromic surveillance is the strengthening of the existing public health infrastructure through practical experience of investigating alarms (Sosin, 2003). The HPA surveillance function has benefited from the development of a multidisciplinary network of experts that analyse, interpret and report NHS Direct surveillance trends. This group has also contributed to much of the research and development reported within this thesis.

Implications of the work

Research

Estimating the burden of disease in specific populations

Only a small proportion of the symptoms/syndromes reported to NHS Direct have been examined in detail (approximately 1 in 6 calls are about respiratory or gastrointestinal syndromes). Others, with relevance to public health, e.g. rash and psychological complaints, may become useful in years to come both for acute surveillance and long term monitoring. Likewise, this thesis has highlighted the relatively low use of NHS Direct by old people but has not examined in detail the spatio-temporal distribution of the half a million calls per annum from this age group. NHS Direct record the ethnicity of each caller but this data field has not been included in the surveillance system or studies reported here. An evaluation of the relative uptake of NHS Direct by ethnic group, and attitudes of these sub-groups to telephone health advice, may highlight inequalities in access to health care advice in the UK and opportunities to address such a gap.

Exanthema

Rash is one syndrome that requires further analytical scrutiny. NHS Direct rash calls account for 4.2% of total NHS Direct calls, varying between 3% and 6% throughout the year. Similarly to respiratory and gastrointestinal calls the largest seasonal variation in rash calls occurs in the under ≤5yr age-group, peaking between May and July. Various diseases have been suggested as significant contributors to this seasonality, with chicken pox, parvovirus infection (‘fifth disease’), and scarlet fever the most likely. Although many cases of these diseases are self-limiting, they exhibit a clear seasonality and can lead to school outbreaks and severe congenital disease in pregnant women. A project is currently underway to estimate the contribution of specific exanthema to rises in rash calls, and interpret the added value of monitoring this syndrome for public health surveillance.

Spatial analyses

The results of studies presented in this thesis, and previous studies, have highlighted social deprivation, age and NHS Direct call centre history as predictors of local NHS Direct usage. In combination these factors do not explain the large spatial variation in call rates (seven fold differences in call rates between PCTs). It is recommended that nationally postcoded NHS Direct data be examined within a multi-variate model to determine the influence of local socio-demographic and health-care structures on call rates. A *Poisson* regression model would be a suitable method as has been used previously to analyse diverse data sources at a spatial level (Green et al., 2006). Such an approach may quantify the main drivers of out-of-hours care in the UK, explain the trade off between NHS Direct and other out-of-hours and emergency services, and help determine the most efficient local out-of-hours care model.

Spatial analysis of fever and vomiting calls suggests SatTScan analyses of other syndromes are worth pursuing. Like influenza, new strains of norovirus emerge (Lopman et al., 2004) that cause significant illness and disruption on a societal level. Further work is needed to explore the spatial diffusion of rises in vomiting calls in children (as well as adults) and establish a timely gold standard for laboratory surveillance of norovirus. This may also help to explain the sharp December peak in calls, possibly caused by an initial upsurge of community norovirus that goes relatively unreported by outbreak and laboratory surveillance systems. Finally, spatial analyses should not be confined to syndromes suggestive of infections. The value of monitoring heat stroke calls, implemented for the UK heatwave plan, may be increased if clusters of heat related calls could be mapped in real time. This has been demonstrated in the US where clusters of ‘heat stroke’ ambulance dispatches are observed in vulnerable communities (Lu et al., 2007).

UK surveillance policy

Influenza surveillance

The UK Department of Health use the RCGP weekly threshold value of influenza-like-illness reports to trigger National Institute of Clinical Excellence (NICE) guidelines on the use of antiviral drugs such as Oseltamivir. Recent discussion of this annual decision has focussed on the need for an informed decision based on many data sources, rather than a single surveillance system (currently the RCGP). It is recommend that the HPA and DH include NHS Direct syndromic surveillance thresholds for cold/flu and fever calls alongside clinical and laboratory data to inform decisions around antiviral prescribing, and for informing the public, media and health professionals of influenza-like-illness levels in the community.

Self sampling for microbiological testing

Although NHS Direct callers have successfully self-sampled for influenza and RSV, this mechanism has not been tested for other pathogens. There are areas where laboratory surveillance data may not accurately describe community morbidity in a timely way, particularly for diseases that may be self-limiting but still highly disruptive on a societal level (e.g. norovirus, childhood exanthema). It is recommended that further pilot studies be conducted to examine the range of gastrointestinal pathogens and exanthema that lead to NHS Direct calls (in children and adults). As well as providing useful baseline data, this would strengthen the NHS Direct self-sampling mechanism should it be needed for more unusual events (e.g. pandemic influenza).

Accident and Emergency department syndromic surveillance

NHS Direct data represent a generally mild spectrum of disease. The service is not universally used unlike GP services and Accident and Emergency departments (A&E). A&E surveillance is undeveloped in England and Wales. It is recommended that the HPA now build on its expertise in syndromic surveillance, and learn from the numerous examples of A&E surveillance in the US (e.g. Heffernan et al., 2004; Dembek et al., 2003), to *pilot* an A&E syndromic surveillance system in the UK. This will help quantify the value of surveillance of unwell individuals' by-passing primary care and reporting directly to A&E.

NHS Direct

Case definitions

An important principle of the NHS Direct syndromic surveillance project has been to cause as little disruption as possible to NHS Direct work patterns. To this end the surveillance system has used only routinely generated call data already collected by NHS Direct for management purposes. There are no formal case definitions within the system, as ICD9 codes, laboratory confirmations, and exposure details are not present within the available data. The structure of the NHS Direct database does not allow for extracting groups of symptoms linked to an individual caller. This functionality could improve the sensitivity of the case definitions and possibly the sensitivity/specificity of surveillance signals. For example, a caller reporting diarrhoea with blood may have E Coli O157 infection. A caller with fever, cough, headache and shivers implicates influenza. This functionality would also allow new syndromes to be extracted in response to a novel or emerging infection. It is recommended that during successive new releases of CAS, efforts are made to enable extraction of multiple symptoms, and that these data are made available to the research and public health communities.

Responsive messaging

When significant trends in NHS Direct syndromic surveillance data emerge, this intelligence is immediately forwarded to staff operating the NHS Direct responsive messaging service (RM). RM is an operational tool designed to help manage service demand and to inform patients of relevant topical information (e.g. 'flu and colds', 'diarrhoea and vomiting'). Syndromic trends may act either as a trigger to initiate messages or to provide statistical support for the ongoing operation of these messages. It is a challenge for the HPA to ensure that sudden surges in calls (e.g. vomiting, fever), that may vary in their seasonal timing, trigger immediate action by NHS Direct to play relevant messages to callers. It is also recommended that NHS Direct explore, in collaboration with academia or the HPA, the value of data derived from the NHS Direct website (http://www.nhsdirect.nhs.uk/) for health protection purposes. Internet based monitoring of ILI has been successful in the Netherlands, Belgium and Portugal (Marquet et al., 2006; van Noort et al., 2007). These systems rely on active participation (emails or internet forms), whereas NHS Direct would be able to utilise routine data about web hits, thus reducing costs and making the surveillance effectively passive in nature.

Surveillance add-on module

The NHS Clinical Assessment System (CAS) is being marketed outside the UK by its provider company Clinical Solutions. It would therefore seem sensible to include with the CAS an add-on surveillance module for potential use in other countries. The format of this module could range from a simple set of data extraction reports suitable for syndromic surveillance (as currently used), to fully integrated surveillance encompassing analysis and display functionality. The ability of the software to map call data in real-time would also prove useful for analysing local demand for telephone triage, and for evaluating the impact of promotional activities.

International

The International Health Regulations (IHR) has recently been strengthened to improve internal health protection arrangements and international cooperation (Baker and Fidler, 2006). Timely surveillance of health data is seen as a powerful tool for implementing the IHR. **Telehealth based surveillance is recommended to other countries where there is a perceived need to develop or enhance real-time monitoring of common syndromes, i.e. where there are recognised shortfalls in existing primary care surveillance.** This syndromic approach may provide particular added value for identifying and monitoring community wide rises gastrointestinal and respiratory infections (e.g. norovirus and influenza). The data may also be used for initiating responsive surveillance during major environmental and climatic incidents; and providing real-time intelligence and reassurance for emergency services, the media, and the public. However, unless call rates are substantially higher than in the UK, it may not be possible to use the data for providing early warning of localised disease outbreaks.

The statistical control chart methodology, used to detect significantly high NHS Direct call levels, employs regression models updated by a statistician every 4 months. The bespoke nature of these analyses, coupled with the unique NHS Direct CAS coding system, means the portability of the surveillance system outside of the UK is low. If data from other telehealth systems are to be used for surveillance, open source syndromic surveillance software (e.g. RODS (RODS, 2007), EARS (EARS, 2007)) are a viable and cost effective option. These systems include statistical algorithms that do not require denominator populations for analyses, instead using total syndrome counts or ratios of syndrome to total counts as indicators. This can be an advantage if the population coverage of telephone triage system is unknown or heterogeneous in nature.

The choice of software platform and statistical technique, however, are not the only important considerations. The processes of epidemiological interpretation and effective dissemination of surveillance alerts will ultimately determine whether the data are used to inform public health action. It is recommended that national institutions undertake this role (e.g. National Institute for Public Health and the Environment (RIVM) in the Netherlands, Institute for Public Health Surveillance (InVS) in France). These institutions are able to take advantage of national public health networks, existing surveillance expertise, and in-house comparison of syndromic data against complementary surveillance sources.

Lessons have been learnt whilst conducting research and operating the NHS Direct syndromic surveillance system on a day-to-day basis. From this comes the following list of recommendations, intended for public health institutions considering using telephone triage data for health protection purposes.

System design

1. Form a multi-disciplinary, dedicated team of analysts and public health personnel.
2. Define and involve the systems stakeholders as early as possible.
3. Set expectations at an appropriate level by describing, if possible, what types of events (infectious or otherwise) the system will and will not detect.
4. Develop analysis and investigation protocols that are linked and integrated into the public health response (i.e. to local public health teams, linking syndromic surveillance with rapid microbiological sampling and near patient tests).
5. Ensure the data cover as wide a population and area as possible and define the system's representativeness.
6. Use telehealth data that are routinely available and automated.
7. If possible, link multiple symptoms rather than single clinical algorithms to individual callers.

System operation

8. Provide clear, useful and simple routine surveillance outputs.
9. Monitor telehealth data alongside data from traditional clinical and laboratory surveillance systems to enhance epidemiological interpretation.
10. Focus on areas where telehealth data can provide a unique picture of community morbidity and ‘added value’ over existing surveillance.
11. Expect software changes that interrupt data flow and affect coding and baseline data.

System evaluation

12. Incorporate, into the surveillance system, a method of evaluating the impact of syndromic surveillance outputs and research on public health practice.
13. Attempt to define what pathogens the telehealth syndromes most likely indicate throughout the year. This will aid interpretation of syndromic trends and make surveillance outputs more persuasive.
14. Syndromic surveillance systems miss more than they detect so focus on the positive predictive value as well as the sensitivity of surveillance ‘signals’.

Summary of conclusions and recommendations

Conclusions

- The NHS Direct surveillance system is broadly representative of primary care usage of those aged below 65 years. Call rates are highest for children under 5 and in areas where social deprivation is just above the national average.
- By using proportions of syndromic calls, rather than rates, it is possible to control for much of the spatial and temporal heterogeneity in call rates.
- The surveillance system has detected acute rises in specific syndromes at regional and national level. This seasonality is determined by viral rather than bacterial infections.
- NHS Direct 'cold/flu' and fever calls rise and peak at the same time as peaks in influenza-like-illness and influenza recorded by clinical and laboratory surveillance systems.
- Numerical threshold values for 'cold/flu' (1.2% of total calls) and fever calls 5-14 yrs (9%) consistently provide 1-2 weeks advance warning of national rises in influenza A and B.
- Spatio-temporal analyses of fever calls for the 5-14yr age group provide a timely and unique description of the spatial evolution of national influenza outbreaks.
- Half the seasonal variation in respiratory calls is estimated to be due to RSV and influenza.
- Self-sampling by NHS Direct callers resulted in a 16% influenza detection rate and is a feasible method of enhancing community based surveillance programmes.
- Nearly half of NHS Direct callers report their gastrointestinal symptoms only to NHS Direct. The surveillance system, therefore, monitors a proportion of illness otherwise unrecorded by surveillance systems.
- Rotavirus and norovirus are estimated to be the most important enteric pathogens determining the seasonality of diarrhoea and vomiting calls to NHS Direct about those under 5 years.
- Retrospective and prospective analyses of call data have not, in any consistent way, demonstrated value for local outbreak detection.
- To date, no deliberate release of either a chemical or biological agent has been detected by the NHS Direct syndromic surveillance system, or any other surveillance system in the UK.
- The main benefits of using NHS Direct call data for surveillance purposes are to provide early warning of regional and national rises in infectious disease (particularly

common childhood infections) and illness caused by environmental factors (e.g. heatstroke, hay fever).

- Additional uses of the data (cited by stakeholders) are for validating other hard and soft intelligence sources, managing winter pressures on the NHS, providing surveillance response and reassurance during major incidents, and for handling the media.

Recommendations

It is recommended that:

- Other countries pursue surveillance of telehealth calls if there is a perceived need for real-time monitoring of common syndromes.
- The HPA and NHS Direct work to:
  - extract data about multiple symptoms reported by a single caller, and
  - explore further spatio-temporal surveillance of syndromes suggestive of influenza, norovirus and heatstroke.
- The UK Department of Health use NHS Direct ‘cold/flu’ and fever thresholds to inform decisions about when to trigger antiviral prescribing for treatment of influenza.
- The HPA build on its expertise in syndromic surveillance to pilot a real-time hospital A&E surveillance system.
- NHS Direct develop in-house software functionality capable of automatically mapping calls.
- An academic research study is conducted to determine the effects of local socio-demographic factors and out-of-hours care models on call rates.
- NHS Direct explore, in collaboration with academia or the HPA, the value of data derived from NHS Direct online for health protection surveillance purposes.

References

Babin SM, Cutchis PN, Tabernero N, et al. A Simple Method of Using Linked Health Data in Syndromic Surveillance. Advances in Disease Surveillance 2007; 4: 83.

Baker M, Smith GE, Cooper DL, et al. Early Warning and NHS Direct: a role in community surveillance? J Public Health 2003; 25:362-368.

Baker MG, Fidler DP. Global public health surveillance under new international health regulations. Emerg Infect Dis 2006;12(7):1058-65.

Balter S, Weiss D, Hanson H, et al. Three years of emergency department gastrointestinal syndromic surveillance in New York City: what have we found? MMWR 2005; 54 Suppl:175-80.

Berger M, Shiau R, Weintraub JM. Review of syndromic surveillance: implications for waterborne disease detection. J Epidemiol Community Health 2006; 60(6):543-50.

Bon F, Fascia P, Dauvergne M, et al. Prevalence of Group A Rotavirus, Human Calicivirus, Astrovirus, and Adenovirus Type 40 and 41 Infections among Children with Acute Gastroenteritis in Dijon, France Journal of Clinical Microbiology 1999; 37(9): 3055-3058.

Bravata DM, McDonald KM, Smith WM, et al. Systematic review: surveillance systems for early detection of bioterrorism − related diseases. Annals of internal medicine 2004; 140(11): 910−22.

Browne I, Cooper DL, Elam G, et al. Evaluation of the usefulness of syndromic surveillance – real time reassurance during major incidents. Health Protection 2007, Warwick University, September 2007.

Buehler JW, Hopkins RS, Overhage JM, et al. Framework for evaluating public health surveillance systems for early detection outbreaks: recommendations from the CDC Working Group. MMWR 2004; 53(RR05): 1-11.

Buehler JW. Review of the 2003 National Syndromic Surveillance Conference − lessons learned and questions to be answered. MMWR 2004: 53(Suppl): 18−22.

Burt J, Hooper R, Jessopp L. The relationship between the use of NHS Direct and deprivation in southeast London: an ecological analysis. J Publ Hlth Med 2003; 25:174-176.

Carlisle R, Groom LM, Avery AJ, Boot D, Earwicker S. Relation of out of hours activity by general practice and accident and emergency services with deprivation in Nottingham: longitudinal survey. *BMJ* 1998; 316: 520 - 523.

Cooke MK, Crofts JP, Joseph CA, et al. Influenza and other respiratory viruses surveillance in the United Kingdom: October 2002 to May 2003. CDR 2004; Suppl: 1-8.

Cooper DL, Hollyoak V, O'Brien SJ, Vaughan M. Do socio-demographic factors affect calls to NHS Direct which are suggestive of infection? A geographical analysis of NHS Direct call data. *Proceedings of PHLS 27th Annual Scientific Conference 2002;* p48.

Cooper DL. Case study: The Use of Tele-health Data for Syndromic surveillance in England and Wales. In: Lombardo JS, Buckeridge DL, eds. Disease Surveillance: A Public Health Informatics Approach. Wiley: New Jersey 2007; Chap 8: 335-365.

de Wit MAS, Koopmans PS, Kortbeek LM, et al**.** Sensor, a Population-based Cohort **Study** on Gastroenteritis in the Netherlands: Incidence and Etiology. **A**m J of Epidemiol 2001; 154(7): 666-674.

Dembek ZF, Carley K, Siniscalchi A, Hadler J. Hospital admissions syndromic surveillance Connecticut. MMWR 2003; Suppl: 50−2.

Department of Health. Heatwave - Plan for England. 2004. http://www.dh.gov.uk/en/Publicationsandstatistics/Lettersandcirculars/Dearcolleagueletters/DH_074608.

Doroshenko A, Cooper D, Smith G, et al. Evaluation of syndromic surveillance based on National Health Service Direct derived data − England and Wales. MMWR 2005; 54(Suppl):117−22.

EARS. Early Aberration Reporting System. http://www.bt.cdc.gov/surveillance/ears/setup.asp

Espino JU, Hogan WR, Wagner MM. Telephone triage: a timely data source for surveillance of influenza-like diseases. AMIA Annu Symp Proc 2003:215-219.

Fleming DM, Charlton JRH. Morbidity and healthcare utilisation of children in households with one adult: comparative observational study. BMJ 1998; 316:1572-1576.

Fleming DM. Weekly returns service of the Royal College of General Practitioners. Commun Dis Public Health 1999; 2:96-100.

Goddard NL, Joseph CA, Watson JM, Zambon M. Epidemiological features of a new strain of the influenza A virus-influenza A (H1N2) circulating in England and its public health implications. Virus Res 2004;103:53-4.

Goddard NL, Joseph CA, Zambon M, et al. Influenza surveillance in England and Wales: October 1999 to May 2000. Commun Dis Public Health. 2000 Dec;3(4):261-6.

Goddard NL, Kyncl J, Watson JM. Appropriateness of thresholds currently used to describe influenza activity in England. Commun Dis Public Health 2003; 6(3):238-45.

Green CG, O Krause D, Wylie JL. Spatial analysis of campylobacter infection in the Canadian province of Manitoba. Int J Health Geogr 2006; 5:2.

Guerrero AC, Shim T, Kemple S, et al. Norovirus Outbreak Detected by Emergency Department Syndromic Surveillance using RedBat. Advances in Disease Surveillane 2007; 4: 165.

Hakre S, Porter WD, Pavlin JA. Survey of Pre-Clinical Behavior- What Do People Do When They Get Sick?. Advances in Disease Surveillane 2007; 4: 250.

Harris JP, Jit M, Cooper D, Edmunds WJ. Evaluating rotavirus vaccination in England and Wales: Part 1 estimating the burden of disease. Vaccine 2007; 25(20): 3962-3970.

Hawker J, Begg N, Blair I, et al. Communicable Disease Control Handbook. Oxford, Blackwell Science Limited, 2001: p21 and 31.

Hawker JI, Olowokure B, Sufi F, et al. Social deprivation and hospital admission for respiratory infection: an ecological study. Respiratory Medicine 2003: 97(11); 1219-1224.

HPA. Health protection Agency. A Surveillance and Response Strategy: Definition, Principles and Overall Strategy. Surveillance Strategy Programme Development Board 2005.

HPA. Health Protection Agency. Health protection in the 21st Century. Chief Executives Office 2005. http://www.hpa.org.uk/publications/2005/burden_disease/default.htm

Heffernan R, Mostashari F, Das D, et al. Syndromic surveillance in public health practice, New York City. Emerg Infect Dis 2004;10(5):858-64.

Henry JV, Magruder S, Snyder M. Comparison of office visit and nurse advice hotline data for syndromic surveillance--Baltimore-Washington, D.C., metropolitan area, 2002. MMWR 2004; 53 Suppl:112-116.

Hippisley-Cox J, Smith S, Smith G, et al. QFLU: new influenza monitoring in UK primary care to support pandemic influenza planning. Euro Surveill 2006 Jun 22;11(6):E060622.4.

Hurt−Mullen KJ, Coberly J. Syndromic surveillance on the epidemiologist's desktop: making sense of much data. MMWR 2005; 54(Suppl):141-146.

Jones RC, Liberatore M, Fernandez JR, Gerber SI. Use of a prospective space-time scan statistic to prioritize shigellosis case investigations in an urban jurisdiction. Public Health Rep. 2006;121(2):133-9.

Knowles E, Munro J, O'cathain A, Nicholl J. Equity of access to health care. Evidence from NHS Direct in the UK. J Telemed Telecare 2006;12(5):262-5.

Leonardi GS, Hajat S, Kovats RS, et al. Syndromic surveillance use to detect the early effects of heat-waves: an analysis of NHS Direct data in England. Social and Preventive Medicine 2006; 51(4): 194-201.

Lopman B, Vennema H, Kohill E, et al. Increase in viral gastroenteritis outbreaks in Europe and epidemic spread of new norovirus variant. Lancet 2004; 363(9410):682-8.

Loveridge P, Cooper DL, Harris J, Grey J, Smith GE. Comparison of Trends in NHS Direct Gastroenteric Syndromes with Norovirus Outbreaks & Laboratory Data 2004 – 2007. Proceedings Health Protection 2007.

Lu J, Metzger K, Cajigal A, Konty K, Matte T. Identifying and Modelling Spatial Patterns of Heat-Related Illness in New York City. Advances in Disease Surveillane 2007; 4: 255.

Marquet RL, Bartelds AI, van Noort SP, et al. Internet-based monitoring of influenza-like illness (ILI) in the general population of the Netherlands during the 2003-2004 influenza season. BMC Public Health 2006 Oct 4;6:242.

Martin C, Tokars J, Rainisch G, et al. Criteria for Prioritizing Statistical Anomalies Identified in BioSense. Advances in Disease Surveillance 2007; 4: 174.

Marx MA, Rodriguez CV, Greenko J, et al. Diarrheal illness detected through syndromic surveillance after a massive power outage: New York City, August 2003. Am J Public Health 2006;96(3):547-53.

Melegaro A, Edmunds WJ, Pebody R, Miller E, George R. The current burden of pneumococcal disease in England and Wales. J Infect 2006; 52(1):37-48.

Metzger KB, Hajat A, Crawford M, Mostashari F. How many illnesses does one emergency department visit represent? Using a population-based telephone survey to estimate the syndromic multiplier. MMWR 2004;53 (Suppl):106-11.

Moore KM, Edgar B, McGuinness M, O'Connor K. Acute Care Alerting and Integrated Public Health Investigation and Response to Syn-dromic Surveillance Alerts: Example of a Local Food Borne Outbreak. Advances in Disease Surveillance 2007; 2: 113.

Mostashari F, Kulldorff M, Hartman JJ, Miller JR, Kulasekera V. Dead bird clustering: A potential early warning system for West Nile virus activity. Emerging Infectious Diseases 2003; 9:641-646.

Muller-Pebody B, Edmunds WJ, Zambon MC, Gay NJ, Crowcroft NS. Contribution of RSV to bronchitis and pneumonia-associated hospitalisations in English children, April 1995-March 1998. Epidemiol Infect 2002; 129: 99-106.

Munro J, Nicholl J, O'Cathian A, Knowles E. Evaluation of NHS Direct first wave sites. First interim report to the Department of Health. Sheffield: Medical Care Research Unit 1998. University of Sheffield.

Munro J, Sampson F, Nicholl J. The impact of NHS Direct on the demand for out-of-hours primary and emergency care. Br J Gen Pract. 2005 Oct;55(519):790-2.

Murray EL, Heffernan R, Yeung A, Kidoguchi LS, Weiss D. Use of Syndromic Surveillance in the Investigation of Salmonella Wandsworth. Advances in Disease Surveillane 2007; 4: 180.

Nordin JD, Goodman MJ, Kulldorff M, et al. Simulated anthrax attacks and syndromic surveillance. Emerg Infect Dis 2005;11(9):1394-8.

Olowokure B, Hawker J, Weinberg J, Gill N, Sufi F. Deprivation and hospital admission for infectious intestinal diseases. Lancet 1999 Mar 6;353 (9155):807-8.

Olson DR, Paladini M, Mostashari F, et al. Using Age & Syndrome to Characterize Epidemic Winter-Seasonal Acute Gastroenteritis. Advances in Disease Surveillane 2007; 2: 165.

Reingold A. If syndromic surveillance is the answer, what is the question? Biosecur Bioterror 2003;1(2):77-81.

Reis BY, Pagano M, Mandl KD. Using temporal context to improve biosurveillance. Proceedings of the National Academy of Sciences of the United States of America 2003; 100: 1961-1965.

Rodman JS, Frost F, Jakubowski W. Using nurse hot line calls for disease surveillance. Emerg Infect Dis 1998;4(2):329-32.

RODS. The RODS Open Source Project. Open Source Outbreak and Disease Surveillance Software http://openrods.sourceforge.net/.

Rolland E, Moore KM, Robinson VA, McGuinness D. Using Ontario's "Telehealth" health telephone helpline as an early-warning system: a study protocol. BMC Health Serv Res 2006;15;6:10.

Rolland E. Comparing Telehealth and NACRS Influenza Activity to Fluwatch Surveillance Using EARS-X. Advances in Disease Surveillane 2007; 4: 189.

Rowlands S, Moser K. Consultation rates from the general practice research database. Br J Gen Pract 2002; 52:658-660.

Salisbury C. The demand for out-of-hours care from GPs: a review. Fam Pract 2000;17(4):340-7.

Saxena S, Majeed A, Jones M. Socioeconomic differences in childhood consultation rates in general practice in England and Wales: prospective cohort study. *BMJ* 1999; 318: 642-6.

Sosin DM, DeThomasis J. Evaluation challenges for syndromic surveillance making incremental progress. MMWR 2004, vol. 53 (Suppl): 125−9.

Sosin DM. Syndromic Surveillance: The Case for Syndromic Surveillance. Biosec and bioterr 2003: 1(4); 247-253.

St George, I. Cullen, M. Branney, M. The deprivation profile and ethnicity of Healthline callers. New Zealand Family Physician 2006 ; 33(6): 386-389.

Steiner-Sichel L, Greenko J, Heffernan R, Layton M, Weiss D Field investigations of emergency department syndromic surveillance signals-New York City; Centers for Disease Control and Prevention (CDC). MMWR 2004 Sep 24;53 Suppl:184-9.

Turner VF, Bentley PJ, Hodgson SA, et al. Telephone triage in Western Australia. Med J Aust 2002; 176(3):100-3.

van den Wijngaard C, van Asten L, van Pelt W,et al. Syndromic Surveillance for a Large Respiratory Disease Outbreak by Legionella in the Netherlands. Advances in Disease Surveillane 2007; 2: 217.

van Dijk A, McGuinness D, Moore K. Ontario's Telehealth: A Novel Syndromic Surveillance System. Advances in Disease Surveillance 2007; 4: 190.

van Dijk A, McGuinness D, Rolland Ee, Moore KM. Can Telehealth Ontario Respiratory Call Volume be Used as a Proxy for Emergency Department Respiratory Visit Surveillance by Public Health? Advances in Disease Surveillance 2007; 4: 200.

van Noort SP, Muehlen M, Rebelo de Andrade H, Koppeschaar C, Lima Lourenço JM, Gomes MG. Gripenet: an internet-based system to monitor influenza-like illness uniformly across Europe. Euro Surveill 2007 Jul 1;12(7):E5-6.

Yih K, Fuller C, Heisey-Grove D, et al. Evaluating Syndromic Signals from Ambulatory Care Data in Four States. Advances in Disease Surveillane 2007; 4: 210.

Weisskopf MG, Anderson HA, Foldy S, et al. Heat wave morbidity and mortality, Milwaukee, Wis, 1999 vs 1995: an improved response? Am J Public Health. 2002 May;92(5):830-3.

Wheeler JG, Sethi D, Cowden JM, et al. Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. BMJ 1999 Apr 17; 318(7190):1046-50.

Wilson LE, Smith A, Meyer N, et al. Utilisation of Data Derived from a Nurse-Led NHS Access and Information Telephone Helpline (NHS24) - Communicable Disease Management in Scotland. Advances in Disease Surveillance 2006; 2:126.

Yih K, Teates KS, Abrams A, et al. Nurse Call Data for Detection of Influenza-Like Illness. Advances in Disease Surveillance 2006; 1:76.

Summary

Disease surveillance is the collection and analysis health data to provide information for action and to inform decisions relating to public health policy. Surveillance systems in the UK typically rely on data about diagnoses made by clinicians within primary care centers, or laboratory confirmations of specific disease pathogens. In recent years there has been a growth in *syndromic* surveillance systems, particularly in the US. Such systems collect and analyse pre-diagnostic information in a real-time fashion in order to provide early warning of rises in disease, and estimate the health status of the community. This thesis concerns how NHS Direct call data can be used to improve health protection surveillance in England and Wales? NHS Direct is a nurse-led health helpline (telehealth system) providing the population of England and Wales with rapid access to health advice and health information. Health protection surveillance is the surveillance of infectious diseases and environmental hazards that may pose a threat to public health.

General introduction

Chapter 1 is a general introduction. The research question of the thesis is: "What is contribution of NHS Direct data to health protection surveillance in England and Wales?" This is addressed with reference to the fundamental principles of modern surveillance systems; that they have clearly defined aims, are flexible and representative of the population, and achieve an acceptable sensitivity, specificity and positive predictive value. The following studies also evaluate whether routine analyses of NHS Direct call data can be used to provide early warning of significant trends in disease, notably outbreaks of respiratory or gastrointestinal illness. The introductory chapter also provides a setting for the work by describing arrangements for health service provision and disease surveillance in England and Wales, and why NHS Direct data may be potentially useful for syndromic surveillance purposes.

The body of the thesis is divided into three parts. Part one introduces the NHS Direct service, and describes the aims, methodology and representativeness of the NHS Direct syndromic surveillance system. Part two contains analytical studies that:

- compare NHS Direct data against data from laboratory and clinical surveillance systems, and
- examine the utility of the data for outbreak detection and providing early warning of rises in disease.

Part three presents results from the syndromic surveillance system and a validity study to determine whether NHS Direct callers reporting syndromes can also self-sample for laboratory influenza testing.

Part 1: NHS Direct data and syndromic surveillance

Chapter 2 discusses the nature of the NHS Direct call record and issues relevant to the use of the data for disease surveillance and epidemiological purposes.
Clinical decision support software (NHS CAS) is used by NHS Direct to collect callers' demographic details and direct them to the appropriate level of care. Data relating to NHS

Direct calls provide a timely snapshot of symptoms occurring in the community and are summarised in ‘off the shelf’ NHS CAS reports. When interpreting NHS Direct derived data, particular attention should be given to the age distribution of callers, NHS Direct demand surges, call ‘networking’ and changes to NHS CAS clinical algorithms. An increasingly rich source of baseline data, growing body of published work, and a more ‘bedded down’ NHS Direct service is furthering understanding and acceptance of the value of the NHS Direct call record.

The use of Primary Care services in the UK is traditionally high in deprived areas but there has been little research into the effect of deprivation on uptake of NHS Direct. **Chapter 3** addresses the repsentativeness of NHS Direct call data by exploring the impact of deprivation, age and gender on call rates to two NHS Direct sites (West Yorkshire and West Midlands). This ecological study linked 6 months call data to electoral wards and the ‘indices of multiple deprivation 2000’. Using a negative binomial regression model, call rates were analysed by age-group, sex, deprivation level and geographical location. Call rates were highest for children under 5 years with a rate ratio of female to male calls (all ages) of 1.30 (95%CI: 1.27-1.33), this difference being most pronounced in 15-44 year olds ($p<0.001$). There was a four-fold variation in call rates between electoral wards. This work demonstrated that overall demand for NHS Direct was highest in areas where deprivation was at or just above the national average. Additionally, it suggested that the effect of extreme deprivation appears to raise adult call rates but reduce rates about children.

Chapter 4 describes the rationale and methodology of the NHS Direct syndromic surveillance system, and the operational findings of the first two years.

The terrorist activity of 2001 highlighted the need to improve surveillance systems for the early detection of chemical or biological attacks. Building on existing pilot work, the NHS Direct syndromic surveillance system aimed to identify an increase in symptoms indicative of the early stages of illness caused by the deliberate release of a biological or chemical agent, or more common infections. Data relating to 10 key symptoms/syndromes were collected electronically from 23 call centres covering England and Wales and analysed on a daily basis. Statistically significant excesses (‘exceedances’) in calls were automatically highlighted and assessed by a multi-disciplinary team. Between December 2001 and February 2003 there were 1,811 exceedances of which 126 led to further investigation and 16 resulted in ‘alerts’ to local or national health protection teams. Some examples of these investigations are described. Although no deliberate release was detected by the system, it was demonstrated that NHS Direct derived surveillance data complements existing surveillance of common infections.

Part 2: Evaluation studies - external validity and comparison against other data sources

Chapter 5 is a simulation study that tested whether the NHS Direct syndromic surveillance system would detect a local outbreak of cryptosporidiosis. Firstly, data from a historical outbreak of cryptosporidiosis was superimposed onto a statistical model of NHS Direct call data for the North London region. The study modelled whether calls about diarrhoea (a proxy for cryptosporidiosis) exceeded a statistical threshold, thus alerting the surveillance team to the outbreak. On the date that the public health team was first notified during the real outbreak the model results predicted a 4% chance of detection by syndromic surveillance (assuming 5% of cryptosporidiosis cases telephoned NHS Direct). This rose to a 72% chance of detection when it was assumed that 90% of cases telephoned. In conclusion, the NHS Direct surveillance system was unlikely to detect an event similar to the cryptosporidiosis outbreak used for the study. The system was and still is most suited to detecting widespread generalised rises in syndromes in the community.

Currently, clinical influenza incidence thresholds (based on GP diagnoses) are used to help predict the likely impact of influenza and inform health professionals and the public of current activity. **Chapter 6** evaluates the potential of NHS Direct syndromic data for providing early warning of national influenza outbreaks. *Poisson* regression models of clinical and syndromic data were used to derive NHS Direct thresholds. The early warning potential of thresholds was evaluated retrospectively for 2002-2006 and prospectively for winter 2006-2007. Over 5 winters NHS Direct 'cold/flu' and fever calls generally rose and peaked at the same time as clinical and laboratory influenza data. Threshold values for 'cold/flu' (all ages) and fever (5-14 years) calls provided two weeks advanced warning of seasonal influenza activity during three of the four winters studied retrospectively, and 6 days advance warning during prospective evaluation. This led to the recommendation that NHS Direct thresholds are used to inform decisions to trigger antiviral prescribing in the UK, and that age-group specific thresholds be developed for other clinical influenza surveillance systems in the UK and elsewhere.

Chapter 7 explains the seasonal variation in NHS Direct respiratory calls by estimating the contribution of specific respiratory pathogens to this variation. Linear regression models were used to estimate the weekly contribution of specific pathogens to the volume of NHS Direct respiratory calls (all ages and 0-4 years). Annual peaks in NHS Direct cough and difficulty breathing calls occurred in late December, with peaks in 'cold/flu' and fever calls occurring between November and April. It was estimated that respiratory viruses, notably influenza and RSV, were responsible for at least 50% of the seasonal variation in NHS Direct respiratory calls. Rhinovirus and parainfluenza were subsidiary contributors. The main bacterial cause of respiratory calls was *S. pneumoniae*, with an estimated year round contribution, peaking during mid-winter. These results have contributed to national estimates of the burden of respiratory diseases in the UK. They have since enabled the surveillance team us to make pathogen specific interpretation of syndromic data, used to provide early warning of rises in community morbidity.

Influenza and norovirus are common viruses that cause significant disease in the UK and world wide, although data about these viruses are not routinely mapped for surveillance purposes in the UK. **Chapter 8** characterises the geographical origin and diffusion of rises in fever and vomiting calls, proxies for influenza and norovirus infection respectively. Call records about fever (5-14yrs) and vomiting (≥5yrs) were extracted for the period June 2005 to May 2006. A space-time scan statistic was used to retrospectively detect statistically significant clusters of calls on a week by week basis. Two distinct periods of elevated fever calls were identified; the first originated in North West England and spread in a south-east direction, the second began in Central England and moved southwards. The timing and geographical location of these rises in fever calls were similar to a national influenza B outbreak occurring during the same winter. Significantly elevated levels of vomiting calls in South-East England during winter 2005/2006 were also identified, though lack of consistent comparison data make the value of this finding more difficult to assess. The analysis of fever calls provided a timely and unique description of the evolution of a national influenza outbreak, with utility for local and national surveillance.

Part 3: Surveillance results focussing on respiratory and gastrointestinal infections

Chapter 9 describes the age distribution and outcomes of NHS Direct calls suggestive of gastrointestinal (GI) infection, namely diarrhoea, vomiting and food poisoning.
GI calls accounted for 10.3% of total calls (‘diarrhoea’ = 4.9%, ‘vomiting’ = 5.1%) with a high incidence in children under 1 year (23.5% of total calls) and those aged 1-4 years (21.5%). Call outcomes which resulted in further NHS care being recommended accounted for 72.3% of total calls but 54.5% of GI calls. Most gastrointestinal (GI) illness within the United Kingdom goes undetected by routine surveillance so these NHS Direct data provide further insight into GI infection in the community.

Chapter 10 is a descriptive study of the operational results from the surveillance systems second year of influenza surveillance. During the winter of 2000/01, age-group specific data relating to the ‘cold/flu’ algorithm were collected from 6 sentinel NHS Direct sites (population coverage: 16 million). Although this winter was characterised by low influenza activity in the United Kingdom, NHS Direct data demonstrated a peak in the ‘cold/flu’ calls (week 06/01) coinciding with the peak recorded by routine influenza surveillance systems. In addition there was an earlier peak in the proportion of ‘cold/flu’ calls (weeks 52/00 and 01/01) which may have been due to other respiratory infections such as RSV, or increased ‘out of hours’ calls to NHS Direct during the holiday period. It was concluded that the timeliness of NHS Direct data, the uniformity of the NHS Direct clinical support systems, and the potential total population coverage of the service made the call data suitable for surveillance during further winters.

Chapter 11 presents a feasibility study to investigate whether self-sampling by NHS Direct callers could provide viable samples for influenza and RSV culture. Two hundred and ninety four NHS Direct callers were recruited and sent self-sampling kits in the post. Callers were asked to take a swab from each nostril and post them to the laboratory. Forty two percent of the samples were returned, 16.2% were positive on PCR for influenza and 5.6% for RSV. The mean time between the NHS Direct call and laboratory analysis was 7.4 days. These samples provided amongst the earliest influenza reports of the season, detected multiple influenza strains, and augmented the routine national syndromic surveillance system. The implications of this work were that self-sampling was a feasible method of enhancing community based surveillance programmes (syndromic or otherwise) for detection of influenza.

General discussion

Chapter 12 discusses the main findings of this work. The conclusions from other related studies that this author has undertaken or contributed to are also included, namely the impact of high ambient temperatures and NHS Direct calls, and qualitative evaluation of the usefulness of NHS Direct surveillance to users. Recommendations are made for health services research, UK surveillance policy, and an international audience.

NHS Direct data are suitable for surveillance purposes because they are broadly representative of the health seeking behaviour of those aged below 65 years in England and Wales, cover a wide range of syndromes, and are available nationally on a daily basis. The seasonality of syndromes indicative of infection is largely determined by viral rather than bacterial diseases. Half the seasonal variation in respiratory calls is estimated to be due to RSV and influenza. Rotavirus and norovirus are estimated to be the most important enteric pathogens determining the seasonality of diarrhoea and vomiting calls. The surveillance system has detected seasonal trends and acute rises in these syndromes at both regional and national level. However, retrospective analyses and operational results have not, in any consistent way, demonstrated value for local outbreak detection.

The main added value of these data for health protection is for influenza surveillance. This has been demonstrated by retrospective modelling of data and operationally via prospective surveillance. Numerical threshold values for calls about the 'cold/flu' and fever syndromes consistently provide 1-2 weeks advance warning of national rises in influenza A and B. Spatio-temporal analysis of fever calls for the 5-14yr age group also provided a timely and unique description of the spatial evolution of a national influenza outbreak, with utility for local and national surveillance. Additionally, self-sampling by NHS Direct callers is a feasible method of providing laboratory diagnoses, and enhancing syndromic and other community surveillance schemes. Other benefits of using call data for surveillance purposes are in providing early warning of regional and national rises in viral gastroenteritis, and illness caused by environmental triggers (e.g. heatstroke, hay fever).

When interviewed, the system's users (those who receive surveillance alerts) reported that the main benefits of the system were for early warning of rises in community influenza and norovirus infection, validating other hard and soft intelligence sources, managing winter pressures on the NHS, surveillance response and reassurance during major incidents, and handling the media. It is recommended that the HPA explore further the use of the data for prospective geographical surveillance of influenza, norovirus and heatstroke. Also that the HPA and NHS Direct attempt to extract data about multiple symptoms linked to a single NHS Direct caller. Finally, other countries should consider surveillance of telehealth data if there is a perceived need to develop real-time monitoring of common syndromes (i.e. there are recognised shortfalls in existing primary care surveillance).

Acknowledgements

First of all, I would like to thank all those who have contributed to the studies presented in this thesis. In particular I am indebted to my supervisors Jouke van der Zee and Peter Groenewegen; firstly for all their guidance, debate and enthusiasm during this work, and secondly for providing me with the opportunity to study and frequently visit the Netherlands. I am also tremendously grateful to the UK Health Protection Agency for supporting my academic studies, and to my friend, co-author and colleague Gillian Smith for support on every level. Also from the HPA, I thank Carol Joseph for many interesting discussions concerning influenza, Neville Verlander for demystifying medical statistics for me, and Angus Nicoll for championing the syndromic surveillance system and keeping a watching brief over my academic progress. Several others have contributed to papers published in this manuscript and I thank Sarah O'Brien and Vivien Hollyoak for much epidemiological support during my first forays into writing for journals. A debt of gratitude is also extended to Rod Griffiths for supporting this surveillance work around the events of 9/11.

This thesis would not have been possible without the help of Frances Chinemana and Maureen Baker (both formerly of NHS Direct) who were eternally positive in finding solutions to the HPA's surveillance needs. From NHS Direct I am also grateful to Ewan Gerard for contributing to research papers and for being a valuable medical link to NHS Direct. I was fortunate enough to receive high quality data from the staff of NHS Direct, the HPA and the Royal College of General Practitioners Birmingham Research Unit, and I am indebted for this service. Thanks to my father for meticulous and constructive proof reading, to my mother continuing support and interest, and to my friend Nicola Moorhouse for cover design. Last but certainly not least, I will be eternally grateful to my family (Hayley, Mia, Rosa and Ivy) for much patience, love, and letting me hog the family computer on numerous occasions.

About the author

Duncan Cooper was born in London, England in 1968 and grew up in York. After taking his A Level studies he completed a three year Geography Degree at the University of Nottingham (1987-90) specialising in Medical Geography and Coastal Geomorphology. He then worked in a variety of jobs, both at home and abroad, and travelled widely in the USA, Africa and Asia. In 1995 he returned to academia studying for a Masters of Research in the Built Environment at the University of Leeds. This led to his appointment as a Geographical Information Officer at the West Midlands Cancer Intelligence Unit where he worked for three years. In 2000 he began work as a Scientist for the UK Public Health Laboratory Service which became the Health Protection Agency in 2003. Here he was instrumental in developing the NHS Direct syndromic surveillance system and became involved in European and international efforts to develop the utility of syndromic surveillance for public health purposes. He became a PhD candidate with the Netherlands Institute for Health Services Research (NIVEL) in Utrecht in 2006. Since February 2008 he has worked as a Public Health Specialist for Bradford and Airedale Primary Care Trust in Yorkshire.

Summary in Dutch

Surveillance voor Gezondheidsbescherming in Engeland en Wales:

Een analyse van syndroomgegevens van NHS Direct

Samenvatting

Bij het monitoren van ziektes (waarvoor we de ook in de Nederlandse vakliteratuur gangbare term surveillance zullen gebruiken) worden gezondheidsgegevens verzameld en geanalyseerd om informatie te krijgen ten behoeve van acties en beslissingen op het gebied van de volksgezondheid. In Engeland zijn surveillancesystemen altijd gebaseerd geweest op gegevens over diagnoses van clinici in de eerstelijnsgezondheidszorg of bevestiging vanuit het laboratorium van specifieke ziektepathogenen. De laatste tijd is met name in de Verenigde Staten het aantal surveillancesystemen gegroeid dat uitgaat van symptomen en syndromen die eigenlijk voor andere doelen zijn vastgelegd, Met dergelijke systemen wordt prediagnostische informatie, zoals gezondheidsklachten, verzameld en gelijktijdig geanalyseerd om een beeld te krijgen van de gezondheidstoestand van de bevolking en om in een vroeg stadium te kunnen waarschuwen voor een toename in ziektegevallen. In dit proefschrift wordt de vraag beantwoord hoe gegevens betreffende telefonische contacten uit een dergelijk (prediagnostisch) systeem, NHS Direct, gebruikt kunnen worden ter verbetering van de surveillance ten behoeve van gezondheidsbescherming in Engeland en Wales. NHS Direct is een telefonische hulp- en informatielijn, waarmee de bevolking in Engeland en Wales snel toegang heeft tot gezondheidsadviezen en –informatie. Met het surveillancesysteem worden infectieziekten en gevolgen van milieurampen gemonitord die de volksgezondheid kunnen bedreigen.

Algemene inleiding

Hoofdstuk 1 bestaat uit een algemene inleiding. De onderzoeksvraag van het proefschrift is: “Wat dragen de gegevens van NHS Direct bij aan surveillance ten behoeve van gezondheidsbescherming in Engeland en Wales?” Bij het beantwoorden van deze vraag worden de belangrijkste uitgangspunten van moderne surveillancesystemen gevolgd:

- de doelen dienen duidelijk geformuleerd te zijn;
- de systemen moeten flexibel zijn en gegevens opleveren die representatief zijn voor de populatie, en
- zij dienen voldoende gevoeligheid, specificiteit en positieve voorspellende waarde te hebben.

In de verschillende deelonderzoeken in dit proefschrift wordt geëvalueerd of routinematige analyse van de gegevens van NHS Direct gebruikt kunnen worden voor vroegtijdige waarschuwing voor significante veranderingen in ziekte, vooral uitbraken van respiratoire en maag/darmziekten. Dit inleidende hoofdstuk schetst tevens de achtergrond van het onderzoek door een beschrijving van gezondheidszorg en surveillance in Engeland en Wales. Tevens wordt ingegaan op de vraag waarom de gegevens van NHS Direct potentieel bruikbaar zijn voor surveillance op basis van syndroomgegevens.

Dit proefschrift bestaat uit drie delen. In deel één wordt NHS Direct geïntroduceerd; de doelen worden beschreven, evenals de methodologie en de representativiteit van het syndroom surveillancesysteem op basis van NHS Direct.
Deel twee bevat analytische onderzoeken die:

- gegevens van NHS Direct vergelijken met gegevens uit het laboratorium en klinische surveillancesystemen, en
- de bruikbaarheid van de gegevens onderzoeken voor detectie van uitbraken en vroegtijdige waarschuwing bij stijging van aantallen ziektegevallen.

In deel drie worden de resultaten gepresenteerd van het syndroom surveillancesysteem en van een valideringsonderzoek waarbij bellers van NHS Direct die melding maken van syndromen zelf wat slijm afnemen en insturen voor een influenzatest in een laboratorium. Tenslotte worden in het laatste hoofdstuk de belangrijkste resultaten van het proefschrift besproken.

Deel 1: NHS Direct gegevens en surveillance op basis van syndromen

In **Hoofdstuk 2** worden de aard van de vastgelegde gegevens inzake de telefonische contacten met NHS Direct besproken en bovendien onderwerpen die te maken hebben met het gebruik van de gegevens voor surveillance en epidemiologische doeleinden.

NHS Direct maakt gebruik van ondersteunende software voor klinische beslissingen (NHS Clinical Assessment System, kortweg NHS CAS) om de bellers naar het juiste zorgniveau te sturen. Deze software wordt ook gebruikt om demografische gegevens van bellers vast te leggen die vergeleken kunnen worden met de bevolking als geheel. Gegevens op basis van de telefonische contacten met NHS Direct geven een vroegtijdige momentopname van symptomen die zich in de bevolking voordoen en worden in standaardrapportages samengevat. Bij de interpretatie van deze aan NHS Direct ontleende gegevens moet in het bijzonder aandacht worden geschonken aan de leeftijdsverdeling van de bellers, aan schommelingen in het aantal bellers naar NHS Direct, en aan de mogelijkheid dat bij grote drukte bij het ene centrum telefoontjes worden doorgesluisd naar een ander centrum dat op dat moment over meer capaciteit beschikt. Tenslotte moet men bedacht zijn op veranderingen in de klinisch algoritmen die binnen NHS CAS worden gebruikt. Doordat er steeds meer en rijkere gegevens beschikbaar komen, het aantal publicaties op basis van NHS Direct gegevens groeit en NHS Direct ook een steeds normaler onderdeel van de gezondheidszorg wordt, worden de rapportages beter begrepen en geaccepteerd als surveillance-informatie.

In Engeland is het gebruik van eerstelijnszorg in achterstandsgebieden vanouds hoog, maar er is nog maar weinig onderzoek gedaan naar het effect van wonen in een achterstandsgebied op het gebruik van NHS Direct. In **Hoofdstuk 3** wordt de representativiteit van de gegevens over telefonische contacten met NHS Direct onder de loep genomen. Dat wordt gedaan door de invloed van de mate van deprivatie van het woongebied en van leeftijd en geslacht van de bellers op de proportie bellers te onderzoeken in twee NHS Direct werkgebieden (West Yorkshire en West Midlands). Bij dit zogenaamde ecologische onderzoek werden gegevens van telefonische contacten gedurende 6 maanden gekoppeld aan de kiesdistricten ('electoral wards') waar de bellers wonen en aan de 'index meervoudige deprivatie 2000', een in Engeland gebruikelijke maat voor deprivatie. Er werd een negatief binomiaal regressiemodel gebruikt om de proportie bellers te analyseren per leeftijdsgroep, sekse, deprivatieniveau van het kiesdistrict en geografische lokatie. Er werd naar verhouding het meest gebeld voor kinderen onder de 5 jaar. De verhouding van de proportie gesprekken van vrouwen ten opzichte van mannen in alle leeftijdscategorieën was 1.30 (95%CI: 1,27-1,33). Het verschil was het duidelijkst bij 15-44 jarigen. De proportie bellers verschilde een factor vier tussen de kiesdistricten. Dit deelonderzoek liet zien dat het totaal aantal bellers naar NHS Direct het hoogst was in gebieden met een deprivatieniveau gelijk aan of net boven het landelijk gemiddelde. Ook lijkt het erop dat extreme deprivatie samen gaat met een hogere proportie bellers onder volwassenen, maar een lagere voor kinderen.

In **Hoofdstuk 4** worden de achterliggende gedachte en de methodologie van het syndroom surveillancesysteem op basis van NHS Direct besproken, evenals de operationele resultaten van de eerste twee jaar.
De terroristische aanval op het World Trade Center in 2001 gaf voedsel aan de noodzaak surveillancesystemen te verbeteren ten einde aanvallen met chemische of biologische wapens vroegtijdig vast te stellen. Voortbouwend op eerder pilotonderzoek zou het syndroom surveillancesysteem van NHS Direct een toename van symptomen moeten kunnen waarnemen die vroege stadia van ziekte aangeven, bijvoorbeeld veroorzaakt door opzettelijke verspreiding van een biologisch of chemisch agens, of door algemeen voorkomende infecties. Dagelijks werden elektronisch vastgelegde gegevens over tien belangrijke symptomen/syndromen in 23 call centra, verspreid over heel Engeland en Wales, geanalyseerd. Statistisch significante verhogingen in het aantal bellers werden automatisch geselecteerd en door een multidisciplinair team beoordeeld. Tussen december 2001 en februari 2003 was er sprake van 1811 verhogingen waarvan er 126 aanleiding waren voor nader onderzoek en 16 resulteerden in een alarm voor de plaatselijke of landelijke gezondheidsbeschermingsteams. Er worden een paar voorbeelden van deze onderzoeken beschreven. Hoewel er geen opzettelijke verspreiding van agentia door het systeem werden ontdekt, werd wel aangetoond dat de monitorgegevens op basis van NHS Direct een aanvulling zijn op de bestaande surveillancesystemen van algemeen voorkomende infecties.

Deel 2: Evaluatieonderzoeken -- externe validiteit en vergelijking met andere gegevensbronnen

Hoofdstuk 5 doet verslag van een simulatieonderzoek om na te gaan of het syndroom surveillancesysteem op basis van NHS Direct een lokale uitbraak van cryptosporidiosis (diarree veroorzaakt door een parasiet) zou detecteren. Daarbij werden gegevens van een historische 'echte' uitbraak van cryptosporidiosis gelegd op een statistisch model gebaseerd op gegevens uit telefonische contacten van NHS Direct voor de regio noord-Londen. Het onderzoek modelleerde de gegevens om te zien of het aantal telefoontjes over diarree (een symptoom van cryptosporidiosis) een statistische drempel zou overschrijden waardoor het volksgezondheidsteam zou worden gewaarschuwd voor een uitbraak. Op de datum dat het gezondheidsteam voor het eerst werd gealarmeerd tijdens de werkelijke uitbraak, voorspelden de modelgegevens 4% kans op detectie door syndroom surveillance, als men ervan uitging dat 5% van de gevallen van cryptosporidiosis NHS Direct zouden bellen. Dit getal steeg naar 72% detectiekans als werd aangenomen dat 90% van de gevallen zouden bellen. De conclusie was dat het surveillancesysteem op basis van NHS Direct een gebeurtenis zoals de cryptosporidiosis-uitbraak die voor het onderzoek werd gebruikt, waarschijnlijk niet zou detecteren. Het systeem was, en is nog steeds, meer geschikt om wijdverspreide, algemene stijgingen in syndromen in de bevolking te detecteren in plaats van een lokale uitbraak van een specifieke ziekte.

Op dit moment worden drempelwaarden voor de incidentie van influenza, afgeleid uit klinische gegevens (diagnoses van huisartsen), gebruikt om het te verwachten effect van influenza te helpen voorspellen en gezondheidsexperts en het publiek te waarschuwen voor naderende griepepidemieën. In **Hoofdstuk 6** wordt geëvalueerd in hoeverre syndroomgegevens van NHS Direct vroegtijdig kunnen waarschuwen voor nationale uitbraken van influenza. Met behulp van Poisson-regressiemodellen werden klinische en syndroomgegevens geanalyseerd om de drempelwaarden voor NHS Direct vast te stellen. De mogelijkheden om op basis van deze drempelwaarden vroegtijdig te waarschuwen werden retrospectief geëvalueerd voor de jaren 2002-2006 en prospectief voor de winter van 2006-2007. In deze vijf winters steeg en piekte het aantal telefoontjes naar NHS Direct over ‘griep of een koutje’ en koorts over het algemeen gelijk met de klinische en laboratoriumgegevens. De drempelwaarden voor telefonische contacten over ‘griep of een koutje’ (voor alle leeftijden) en koorts (5-14 jaar) leidden twee weken eerder tot waarschuwingen voor seizoensgebonden influenza-activiteit voor drie van de vier onderzochte winters uit het verleden. Bij de voorspellende evaluatie gaven deze drempelwaarden zes dagen eerder een waarschuwing. Op grond hiervan is de aanbeveling gedaan NHS Direct drempelwaarden te gebruiken ter ondersteuning van beslissingen om te beginnen met het voorschrijven van antivirale middelen. Daarnaast is aanbevolen om ook voor bestaande klinische influenza surveillancesystemen in Engeland en elders leeftijdsspecifieke drempelwaarden te ontwikkelen.

In **Hoofdstuk 7** staat de seizoensvariatie in telefonische contacten met NHS Direct naar aanleiding van ademhalingsklachten centraal. In dit deelonderzoek wordt de bijdrage van specifieke respiratoire pathogenen aan deze variatie geschat. Lineaire regressiemodellen werden gebruikt om van week tot week de bijdrage van specifieke pathogenen aan het aantal telefonische contacten (voor kinderen van 0-4 jaar) met NHS Direct te schatten. Telefonische contacten met NHS Direct voor 0-4 jarigen over ‘hoesten’ en ‘moeite met ademhalen’ vertonen traditioneel eind december een piek. Pieken in het aantal meldingen van ‘griep of een koutje’ en ‘koorts’ traden daarentegen vooral op tussen november en april. Waarschijnlijk waren respiratoire virussen, in het bijzonder influenza en respiratoir syncitieel virus (RSV), verantwoordelijk voor ten minste de helft van de seizoensvariatie in telefonische contacten met NHS Direct naar aanleiding van respiratoire aandoeningen. Daarnaast spelen rhinovirus en parainfluenza hierbij een rol. De belangrijkste bacteriële oorzaak in het geval van telefonische contacten naar aanleiding van respiratoire aandoeningen was *S. pneumoniae*, die naar schatting het hele jaar bijdroeg, met een piek midden in de winter. Deze resultaten zijn op hun beurt gebruikt voor landelijke schattingen van de ernst van respiratoire aandoeningen in Engeland. Hierdoor is het mogelijk geworden om pathogeenspecifieke interpretaties te geven van de syndroomgegevens die worden gebruikt om vroegtijdig te waarschuwen voor stijgingen in morbiditeit in de bevolking.

Influenza en norovirus zijn algemeen voorkomende virussen en belangrijke ziekteveroorzakers in Engeland en wereldwijd. Gegevens over deze virussen worden in Engeland echter niet routinematig in kaart gebracht voor surveillance doeleinden. In **Hoofdstuk 8** wordt de geografische oorsprong en de verspreiding van stijgingen in aantallen telefonische contacten naar aanleiding van 'koorts' en 'overgeven' gekenschetst. Koorts en overgeven zijn namelijk indicatief voor respectievelijk influenza en norovirusinfectie. Daarvoor werden registratiegegevens van telefonische contacten over koorts (voor kinderen van 5-14 jaar) en overgeven (≥5 jaar) gebruikt voor de periode juni 2005 tot mei 2006. Met behulp van een statistische model om afwijkingen in ruimte en tijd vast te stellen werden retrospectief statistisch significante clusters van telefonische contacten met NHS Direct van week tot week geanalyseerd. Er werden twee afzonderlijke perioden van verhoogd aantal telefonische contacten over koorts waargenomen; één begon in Noordwest Engeland en verspreidde zich in zuidoostelijke richting, de tweede begon in Centraal Engeland en bewoog zich zuidwaards. De timing en geografische locatie van deze stijgingen in koortsmeldingen kwamen overeen met die van een nationale uitbraak van influenza B gedurende dezelfde winter. Ook werden significant verhoogde niveaus van contacten inzake overgeven in Zuidoost Engeland gevonden in de winter van 2005/2006. De waarde van deze bevinding is moeilijker te beoordelen door een gebrek aan bruikbare klinische of labgegevens. De analyse van de NHS Direct telefoontjes over koorts resulteerde in een vroegtijdige en unieke beschrijving van de diffusie van een uitbraak van influenza over het land.

Deel 3: Resultaten van surveillance ge richt op respiratoire en gastrointestinale infecties

In **Hoofdstuk 9** worden telefoontjes met NHS Direct beschreven naar aanleiding van klachten die wijzen in de richting van gastrointestinale infectie (GI), namelijk diarree, overgeven en voedselvergiftiging. Telefonische contacten die wijzen op GI vormden 10% van het totaal aantal telefoontjes uit ('diarree' = 4,9%, 'overgeven' = 5,1%) met vooral een hoge incidentie bij kinderen onder de 1 jaar (23,5% van het totaal aantal telefoontjes) en kinderen van 1-4 jaar (21,5%). Telefonische contacten met NHS Direct die resulteerden in doorverwijzing voor verdere NHS-zorg, maakten 72,3% van het totaal aantal contacten uit, maar 54,5% van de GI contacten. De meeste gastrointestinale (GI) aandoeningen binnen Groot-Brittannië worden niet opgemerkt met andere routinematige surveillance. Deze gegevens op basis van NHS Direct geven dus meer inzicht in het optreden van gastrointestinale infecties in de bevolking dan bestaande systemen.

Hoofdstuk 10 bestaat uit een beschrijvend onderzoek van de operationele resultaten van het tweede jaar van de influenzasurveillance. In de winter van 2000/01 werden leeftijdsspecifieke gegevens verzameld met betrekking tot klachten rond ‘griep of een koutje’. De gegevens betroffen zes NHS Direct regio’s (met een totale bevolking van 16 miljoen). Hoewel de winter van 2000/01 gekenmerkt werd door een lage influenza-activiteit in Engeland, vertoonden de gegevens van NHS Direct een piek in de telefoontjes over ‘griep of een koutje’ (in week 6 van 2001) die samenviel met de piek in klinische influenza surveillancesystemen. Bovendien was er daarvoor ook al een piek in het aandeel telefoontjes over ‘griep of een koutje’ (in de laatste week van 2000 en de eerste week van 2001). Dit kan te maken hebben gehad met andere respiratoire infecties, zoals RSV, of met een hoger dan normaal beroep op NHS Direct doordat veel gezondheidszorginstelling in de vakantieperiode gesloten waren. De conclusie was dat de meldingsgegevens geschikt waren voor surveillance in de komende winters, omdat de gegevens op basis van NHS Direct zowel snel beschikbaar zijn, als gebaseerd zijn op uniform toegepaste klinische algoritmen binnen NHS Direct en bovendien de totale bevolking dekken.

In **Hoofdstuk 11** wordt een haalbaarheidsonderzoek beschreven waarin wordt onderzocht of ‘self-sampling’ door NHS Direct bellers bruikbaar materiaal opleveren voor influenza- en RSV-labbepalingen. Er werden 294 NHS bellers in principe bereid gevonden mee te doen en hen werd een self-sampling pakket per post toegezonden. Bellers werd gevraagd een uitstrijkje uit ieder neusgat te nemen en dit naar het laboratorium te sturen. Tweeënveertig procent van de uitgezette pakketten werd teruggezonden, 16,2% was positief op PCR (polimere chain reactie, een techniek om virussen te identificeren) voor influenza en 5,6%voor RSV. De gemiddelde tijd tussen de melding bij NHS Direct en de laboratoriumanalyse was 7,4 dagen. Deze uitstrijkjes die helemaal aan het begin van het influenza seizoen werden genomen, toonden meerdere influenzastammen aan en vormden een goede aanvulling op het bestaande syndroom surveillancesysteem. De implicatie van deze studie is dat self-sampling een haalbare aanvulling is op de gangbare influenzasurveillance.

Algemene discussie

In Hoofdstuk 12 worden de belangrijkste bevindingen op een rij gezet. De conclusies van andere onderzoeken van deze auteur, of waar deze auteur een bijdrage aan heeft geleverd, worden ook genoemd. Deze gaan enerzijds over de impact van hoge buitentemperaturen op het aantal telefonische contacten met NHS Direct, en anderzijds over een kwalitatieve evaluatie van het nut van surveillance via NHS Direct vanuit het gezichtspunt van gebruikers van de resultaten. Er worden aanbevelingen gedaan voor verder gezondheidszorgonderzoek en over het surveillancebeleid in Engeland. Ten slotte worden de internationale implicaties besproken.

Gegevens van NHS Direct zijn geschikt voor surveillance omdat zij ten eerste in grote lijnen representatief zijn voor de manier waarop mensen onder de 65 in Engeland en Wales een oplossing zoeken voor hun gezondheidsproblemen, ten tweede een uitgebreide reeks syndromen dekken, en ten derde van dag tot dag beschikbaar zijn. De seizoensgebondenheid van syndromen die op infecties wijzen wordt in grote lijnen eerder bepaald door virale dan door bacteriële aandoeningen. Naar schatting is de helft van de seizoensgebonden variatie in contacten met NHS Direct inzake respiratoire aandoeningen toe te schrijven aan RSV en influenza. Rotavirus en norovirus zijn, zo vermoeden we, de belangrijkste spijsverteringspathogenen die de seizoensgebondenheid van de telefoontjes over diarree en overgeven bepalen. Met het surveillancesysteem zijn trends over de seizoenen vastgesteld, evenals acute stijgingen in deze syndromen, zowel regionaal als nationaal. Echter, lokale uitbraken zijn op basis van de retrospectieve analyses en de operationele resultaten niet consistent vast te stellen.

De belangrijkste toegevoegde waarde van de NHS Direct gegevens voor gezondheidsbescherming betreft de influenzasurveillance. Dit is aangetoond met retrospectieve modellering van gegevens en operationeel via prospectieve surveillance. Numerieke drempelwaarden voor telefonische contacten over ‘griep of een koutje’ en koortssyndromen geven consistent één tot twee weken eerder waarschuwingssignalen voor een landelijk toename van influenza A en B. Tijd-ruimtelijke analyse van koortstelefoontjes voor 5-14 jarigen leverde ook een tijdige en unieke beschrijving van de verspreiding van influenza, bruikbaar voor lokale en landelijke monitoring. Self-sampling door bellers van NHS Direct is een uitvoerbare methode om snel laboratoriumdiagnoses te verkrijgen en toe te voegen aan het arsenaal aan syndroom surveillance en andere systemen van volksgezondheidsurveillance. Bijkomend voordeel van het gebruik van NHS Direct gegevens voor surveillancedoelen is dat zij vroegtijdig waarschuwen voor regionale en landelijke stijgingen in virale gastro-enteritis, en ziekten veroorzaakt door omgevingsfactoren (bijvoorbeeld zonnesteek, hooikoorts).

De gebruikers van het systeem (zij die waarschuwingssignalen uit het surveillancesysteem ontvangen) gaven in interviews aan dat de belangrijkste voordelen van het systeem zijn: vroegtijdig waarschuwing voor stijging in influenza en norovirusinfectie; validering van andere harde en zachtere gegevens; beter management van het beslag op de NHS gedurende de winter; de mogelijkheid van monitoring en geruststelling van de bevolking bij grote incidenten; en dat het bovendien helpt met de media om te gaan in zulke situaties. Een aanbeveling voor de Health Protection Agency en NHS Direct is om het systeem zo aan te passen dat gegevens geselecteerd kunnen worden over meerdere symptomen die op één en dezelfde beller van NHS Direct betrekking hebben. Ook zou de HPA het gebruik van gegevens ten behoeve van prospectieve geografische surveillance van influenza, door norovirus veroorzaakte darminfecties en zonnesteek verder moeten onderzoeken. Ten slotte, zouden andere landen in overweging kunnen nemen tele-gezondheidsgegevens als surveillance-instrument te gebruiken om te voorzien in real-time surveillance van algemeen voorkomende syndromen, zeker wanneer bestaande surveillancesystemen, bijvoorbeeld op basis van eerstelijnszorg, niet voldoende zijn.