What drives demand for unscheduled care services in Wales?

Authors: Dr Hugo van Woerden and Stephen Williams on behalf of the members of the Unscheduled Care Working Group (Dr Ciaran Humphreys, Dr Mark Temple, Professor Stephen Monaghan, Dr John Watkins, Professor Rob Atenstaedt)

Executive Lead: Professor Peter Bradley, Director of Public Health Development

Date: 19 February 2015

Version: 1a

Publication/ Distribution: Public (Internet)

Summary and purpose:

This report describes a wide range of important ‘drivers’ which contribute towards the growing gap between demand and supply in unscheduled care. It follows on from a 2013 public health report, “External factors (drivers) affecting long term trends and recent pressures on unscheduled care use and performance in Wales” and is one of a number of public health reports developed to support the redesign of unscheduled services in Wales.

This report identifies a range of environmental, structural and behavioural factors that contribute towards rising demand for unscheduled care in Wales and contribute towards the growing gap between supply and demand.

The unscheduled care system in Wales is complex, with each driver interacting with other drivers in the system, potentially amplifying the effect of other drivers. It is difficult to determine the relative proportion of avoidable demand attributable by each driver. In terms of short term demand drivers the biggest challenge is the presence of significant fluctuation in unscheduled care demand and a related mismatch between the resources deployed at any given time and the demand at that time. This is an especially significant issue in Emergency Departments. Weather patterns, major social events and seasonal effects can cause significant
disruption if systems are run at close to capacity. Short term flexing of capacity may be the most effective way of responding to the urgent need to create the ‘headroom’ required to manage fluctuations in unscheduled care demand.

The largest challenge in addressing unscheduled care in the longer term is related to structural issues in the healthcare systems which can be described as ‘failure demand’, that is, “demand caused by a failure to do something or do something right for the customer”. The current system inadvertently pushes patients from primary care and pulls them towards secondary care, encouraging demand in the wrong part of the care pathway.

Patients’ beliefs and behaviours are important and can be partially modified by educational approaches such as the “Choose Well” campaign, but are probably best addressed in the longer term by making options other than Emergency Departments more attractive to patients.

It is important not to lose sight of societal factors such as fuel poverty and an aging population with high levels of co-morbidity as underlying drivers of demand. Responding to these social factors will require combined action by health and social care services and a move towards a greater proportion of care being delivered by non-statutory services.

**Intended audience:**
Unscheduled Care Lead for Wales
Unscheduled Care Board for Wales
NHS staff
1 Introduction

This report summarises the demand drivers that contribute towards the growing gap between demand and supply in unscheduled care. The views expressed in this report are based on a pragmatic review of the relevant literature over the last five years, previous public health reports on unscheduled care, the views of participants in an unscheduled care workshop held in October 2014 and the views of a wider advisory group. This report develops some of the themes that were presented in a previous report, “External Factors (‘Drivers’) Affecting Long-Term Trends and Recent ‘Pressures’ on Unscheduled Care Use and Performance in Wales”\textsuperscript{1}.

Hospitals and health boards in Wales are under significant pressure in relation to unscheduled care. This pressure has often been attributed to the ‘acuity’ of the patients that are presenting to the healthcare system without defining what is meant. This report defines unscheduled care and then explores the drivers of demand under the following headings: short term external drivers that produce fluctuation in demand, long term external drivers of demand, supply side factors that affect the gap between demand and supply, and patient-related behavioural factors.

1.1 Defining unscheduled care

A Royal College of General Practitioners (RCGP) report\textsuperscript{2} states that a range of overlapping terms have been used in relation to unscheduled care, with terms such as “unscheduled care”, “unplanned care”, “emergency care” and “urgent care” sometimes used interchangeably. The following definitions are recommended by the RCGP report:

- **Emergency care** is an immediate response to a time critical health care need
- **Unscheduled care** involves services that are available for the public to access without prior arrangement where there is an urgent actual or perceived need for intervention by a health or social care professional
- **Urgent care** is the response before the next In–Hours or routine (primary care) service is available


\textsuperscript{2} RCGP (2011) Urgent and Emergency Care Clinical Audit Toolkit. London: Royal College of General Practitioners.
• **Urgent and emergency care** is the range of healthcare services available to people who need clinical advice, diagnosis and/or treatment quickly and unexpectedly

A broader definition that has been utilised by Healthcare for London is "Unscheduled care is any unplanned contact with the NHS by a person requiring or seeking help, care or advice. It follows that such demand can occur at any time and that services must be available to meet this demand 24 hours a day. Unscheduled care includes urgent care and emergency care"\(^3\). This definition is particularly useful in the context of primary care.

At a systems level, scheduled and unscheduled care needs to be delivered in a seamless manner, as changes in one part of the system affects the other. It is essential that we remember that all these constructs are ultimately artificial and are at best partial representations of the real world.

A summary of the main drivers of unscheduled care is provided below in relation to short term and long term external drivers, ‘failure demand’ and behavioural factors in service users.

## 2 Short term factors affecting demand

### 2.1 Seasonal factors

Hot and cold weather are both associated with increased demand for unscheduled care. Respiratory illnesses have a distinct seasonal pattern, with an increase in winter largely due to influenza infection leading to hospital admission and excess winter mortality. Other viral infections, such as noro-virus, are more common in summer\(^4\). Both viruses can place significant short term strain on unscheduled care services. Heat waves are associated with increased mortality. However, the maritime climate of the UK limits the importance of this risk factor.

#### 2.1.1 Respiratory infections

Seasonal influenza and other respiratory virus infections can significantly affect demand for unscheduled care, particularly in the winter months. Fortunately, in recent winters seasonal influenza has not reached the “higher than average activity” threshold but primary and secondary care

---


systems still need to ensure that they have the surge capacity to respond to such surges in demand as they are likely to occur every few years. The wide range in the rate of consultations for these infections from year to year is demonstrated in figure 1.

**Figure 1: GP consultation rates for seasonal influenza, Wales**

![Figure 1](image)

**2.1.2 Seasonal patterns in myocardial infarction, stroke and falls**

There is also a clear association between different medical conditions and temperature. In one study, each one degree centigrade reduction in daily mean temperature was associated with a 2.0% cumulative increase in the risk of myocardial infarction over the following 28 days. Drops in temperature also have effects on the incidence of stroke, and falls leading to a range of fractures.

---


2.2 **Staff/demand resource mismatch in daily and weekly cycles**

Across the health and social care system there is evidence of a mismatch of the resources required to meet peaks in demand that are associated with temporal patterns, e.g. the 24 hour pattern, the weekly pattern and the seasonal pattern. This is illustrated below in the case of primary care. Figure 2 indicates that many GP practices appear not to adjust staffing levels to address the regular surge in demand that occurs on a Monday morning. The figure displays the number of practices that had sufficient staff to meet peak demand on Monday mornings without developing lengthy queues. Many GP practices did not have enough staff to meet the surge in demand. The mismatch between demand and staffing levels appears to be replicated across Ambulance services, Emergency Departments, ITUs, theatres, community and social services.

**Figure 2: Number of practices with adequate staff to meet Monday morning calls**

![Graph showing number of practices with adequate staff to meet Monday morning calls](image)

Over the last few decades, employment patterns have made the health service workforce much less flexible\textsuperscript{8,9}. Rotas do not appear to map to daily, weekly and annual patterns of unscheduled care demand\textsuperscript{10}. There is perhaps a need to look at methods that would allow staff to be pulled in to provide cover at short notice (e.g. one to two hours) to address surges in demand\textsuperscript{11}. Based on mathematical modelling, this surge capacity might need to be as high as 50\% of staffing levels in some units.

There could also be better planning of the deployment of resources in relation to major national social events, which are likely to trigger a surge in unscheduled care demand\textsuperscript{12}.

3 Long term external demand drivers

3.1 The aging population

An important long term driver of unscheduled care demand is the aging of the population, accompanied by increasing co-morbidity, medicalisation, frailty and social isolation. As people live longer but have fewer children, there is an increased proportion of the population who are dependent on care in relation to those who can provide care. On average, older people have lower baseline function, greater frailty and lower resilience. This leads to greater need for support for the activities of daily living, tipping over into acute ill health at a lower threshold, and slower recovery from illness, all of which places increased demand on health and social care services.

There has been an increase in the population of Wales as a whole, which has put pressure on some areas. The population of Wales grew by

\begin{thebibliography}{9}
\end{thebibliography}
209,000 between 1991 and 2013\textsuperscript{13}. The population growth is not evenly distributed and therefore has more effect in some areas that on others.

This challenge is particularly acute in Wales. Figure 3 indicates that in a series of international comparisons published by the Welsh Government, Wales had the third highest population aged 80 years and over. This is important because of the impact of this age group on demand for unscheduled care.

**Figure 3: Estimated proportion of population aged 80 years and over, 2009\textsuperscript{14}**


The non-linear nature of frailty with increasing age is shown in figure 4, which is taken from a systematic review of the prevalence of frailty.

**Figure 4: Prevalence of frailty in the community (with 95% CIs)**

The prevalence of frailty in those aged over 85 years is around 26%, whereas the prevalence of frailty in those aged 65-69 years is around 4%. Frailty was defined in this study using a range of factors including strength and agility combined in a validated scoring system.

Frailty seems to be associated with reduced resilience, or reserve, in the face of acute illness. For example, the health of the average 95 year old will take a greater ‘dip’ as a result of a urinary tract infection, as compared to the average 50 year old.

---

3.2 Complex co-morbidity and disease treatment interactions in the elderly

The problem of the aging population is made worse by the presence of increasing co-morbidity with age. The rate of co-morbidity with increasing age is shown in figure 5.

Figure 5: Number of chronic disorders by age group\textsuperscript{16}

![Figure 5: Number of chronic disorders by age group](image)

Most disease guidelines have been designed to optimise the treatment of individuals with only one disease. However, treatment for several conditions leads to disease-disease, disease-treatment and treatment-treatment interactions that are not addressed by guidelines. In the context of unscheduled care, the increasing complexity of treatments affects both demand and supply factors. Wales has a particularly high rate of dispensing of medication (see Figure 6) and potentially an issue with poly-pharmacy.

There is evidence to suggest that these factors are associated with increased unscheduled care demand on healthcare services.

### 3.2.1 Increasing consultation rates

There is also evidence that after adjusting for age, the level of healthcare received by the elderly has risen. As shown in Figure 7, there has been a rise in the average number of primary care consultations per patient per year, particularly in those aged over 75 years. This data was gathered from a large set of GP practices in the UK. It may be that, older people are taking more medication than was the case in the past and so require more intensive follow up by GPs.

---

This graph provides some evidence that there has been increasing ‘medicalisation’ of the whole of society but particularly of old age, with high levels of poly-pharmacy and the initiation of treatment in old age for conditions that would historically have been conservatively managed by observation and palliation rather than aggressive intervention. The combination of the aging population, increased co-morbidity and the ‘medicalisation’ of the elderly population means that Wales faces a major health challenge in relation to healthcare provision for its elderly population. This issue is picked up later in this document, in relation to supply side factors such as the number of beds per head of population in Wales.

3.2.2 Changes in non-NHS service provision

A range of community services are under pressure and providing less support in the community, leading to backward pressure on the discharge of patients from high intensity inpatient care. There has been pressure on social services budgets over a number of years with changes in the threshold at which individuals receive access to support. In some demographic groups there are incentives to look after frail elderly at home.

---

to avoid the high cost of residential or nursing home costs whether this is borne by the individual or family. In other demographic groups there are incentives for carers and families to opt for residential or nursing home care to reduce the cost to the individual or family. The timescales at which assessment progresses are widely recognised as a factor contributing to delayed discharges from hospital care.

A recent exploratory analysis undertaken in conjunction with colleagues in Welsh Government suggests that there has been a fall in private sector residential and nursing home beds in Wales\(^\text{19}\). Some areas are also reporting a reduction in home care packages. These factors reduce the overall pool of resource available and contribute to increased backward pressure on NHS inpatient services. Fewer resources are available in the system as a whole to meet the increasing needs of an aging population.

### 3.3 Fuel poverty and excess winter deaths

Wales has a relatively poor housing stock. A relatively high percentage of properties in deprived areas are in the private sector and are owner occupied, but much of this housing stock was built around 100 years ago and many homes are not insulated to a level that would allow low cost heating and benefit the health of older residents. This has an effect on excess winter mortality, as is shown in table 1.

**Table 1: Excess Winter Mortality, 2012-13**

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Excess Winter Deaths</th>
<th>Excess Winter Mortality Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–64</td>
<td>200</td>
<td>14.5</td>
</tr>
<tr>
<td>65–74</td>
<td>200</td>
<td>10.7</td>
</tr>
<tr>
<td>75–84</td>
<td>400</td>
<td>14.7</td>
</tr>
<tr>
<td>85+</td>
<td>1,100</td>
<td>27.7</td>
</tr>
<tr>
<td>All ages</td>
<td>1,900</td>
<td>18.8</td>
</tr>
</tbody>
</table>

Table 1 indicates that excess winter deaths rise with age. There is also a greater contribution from some diseases as opposed to others. The rise in Excess Winter Deaths is particularly marked as age rises in the case of respiratory disease and injuries, as is shown in table 2.

\(^{19}\) Health statistics Wales 2014, Summary results, Table 15.2 and Chapter 16, Table 16.1, Welsh Government.
### Table 2: Excess Winter Mortality by disease category, 2011-12

<table>
<thead>
<tr>
<th>Age group</th>
<th>Circulatory</th>
<th>Respiratory</th>
<th>Injury</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–64</td>
<td>9.9</td>
<td>35.2</td>
<td>10.6</td>
</tr>
<tr>
<td>65–74</td>
<td>12.9</td>
<td>30.3</td>
<td>18.0</td>
</tr>
<tr>
<td>75–84</td>
<td>15.0</td>
<td>37.3</td>
<td>18.2</td>
</tr>
<tr>
<td>85+</td>
<td>19.9</td>
<td><strong>44.8</strong></td>
<td><strong>24.9</strong></td>
</tr>
<tr>
<td>All ages</td>
<td>16.2</td>
<td>39.7</td>
<td>15.6</td>
</tr>
</tbody>
</table>

Although outside the immediate remit of healthcare services, fuel poverty is a significant long term driver in an aging population in relation to unscheduled care demand.

### 3.4 The effect of socio-economic deprivation

Utilisation of healthcare services is more related to biological age than to chronological age\(^{20}\). Premature aging, years lived with disability are associated with socio-economic deprivation\(^{21}\). There is a need for the level of the provision and funding of unscheduled services to take such socio-economic deprivation into account.

### 4 Supply side factors including ‘failure demand’

Part of the demand for health services arises from the unintended consequences of health care services, which can generate demand by failing to optimise access to care at the appropriate point in a given care pathway. A range of supply side factors contributing to demand have been called ‘failure demand’ by Seddon\(^{22}\). The remainder of this section of the report highlights a number of areas in which an element of failure demand may be operating in the Welsh healthcare system. Other parts of the UK have similar problems. There is evidence of failure demand for example arising from the incentives generated by the payment by results approach in England, which does not affect the Welsh healthcare service to the same extent\(^{23}\).

---


4.1 The push from primary care and the pull towards secondary care

There are structural factors that exert both a push away from primary care services and a pull towards secondary care. These factors have gradually accumulated as a result of inadvertent effects associated with policy decisions that have been taken over a number of decades. Hospital care has been seen as more expert care and therefore treated preferentially within the health economy as a whole. Writing in relation to healthcare in London, Lord Darzi pointed out that the current hospital infrastructure gives the impression that higher quality resources can best be accessed via hospitals. Investment in secondary care has outstripped investment in primary and community care.

The gatekeeper role of GPs has been denigrated and is sometimes seen as something to be circumvented by those who want the best care. GPs no longer have admission rights to hospitals and have poorer access to specialist tests.

A proportion of GP time has been diverted to health improvement work via the introduction of the Quality and Outcomes Framework (QOF) which, though no doubt important, has reduced the capacity for responding to unscheduled care needs by primary care services. The current primary care payment system favours chronic disease management and health promotion work rather than unscheduled care.

The split in day time and Out-of-Hours GP care has had an effect. Many younger GPs rarely take part in Out-of-Hours care and have become de-skilled in relation to urgent and unscheduled care. This has made it increasingly difficult to fill Out-of-Hours GP shifts which are seen as poorly rewarded, high risk activities that have been culturally devalued in the context of primary care as a whole.

The 1990 GP Contract also reduced continuity of care and introduced a range of unintended consequences. The following example illustrates one unintended effect:

A high proportion of unscheduled care demand in primary care is best treated by reassurance. In the late afternoon or evening anxiety levels in carers and family members naturally rises. The perception by patients

---


and carers of the severity of an illness increases as the day progresses as carers become increasingly tired and concerned regarding the prospect of providing care overnight. As a result, phone calls looking for reassurance are often received by GP practices late in the afternoon when no appointments are free. Prior to 2003, when GPs had 24-hour responsibility for patients, there was a strong case for fitting in patients who primarily needed reassurance and who phoned the surgery at 6 pm, because if inadequate reassurance was provided at that stage, a visit would probably be required late in the evening. The introduction of separate Out-of-Hours services has resulted in a reluctance to meet this demand as it is not based on a strict interpretation of ‘clinical need’, as the consequences of inaction falls on the Out-of-Hours service instead. However, reassurance by an Out-of-Hours clinician is less effective than that provided by a known and trusted doctor. As a result a proportion of Out-of-Hours calls that could be treated by reassurance are treated instead by referral to hospital. Hospitals are seen as providing better reassurance, as Emergency Departments often take a more interventionist approach to issues that might have been adequately addressed by simple reassurance from a GP who knew the patient well. This cycle of clinical behaviour leads to ‘failure demand’ with the increased utilisation of resources by Out-of-Hours services and Emergency Departments.

There has been repeated structural change in the NHS, which the public has not caught up with, which also leads to increased inappropriate demand. For example, Health Board names and boundaries have changed. The location of Out-of-Hours services has changed over a number of years. There is a lack of clarity of the role of the NHS Direct and 111 services. In that context, Accident and Emergency departments have been a stable concept in the public’s consciousness. Patients are not knowledgeable about where Out-of-Hours services are located, and they can be difficult to find in the middle of the night, whereas the public is very aware where their nearest Emergency Department is because road signs leading to them are highlighted in red.

The above example is not an isolated one. The unintended consequence of policy decisions is a major healthcare problem which has been called the ‘Whitehall Effect’. Such structural issues are difficult to address, but do point to a need to placing a greater value on continuity of care, high quality Out-of-Hours care and the use of resources to make Out-of-Hours care attractive to high quality staff.

---

27 Davies P. (2014). Exposing the myth that reconfiguration of clinical services improves quality and saves money. BMJ 349, g7172.

4.2 Medicalisation of society, increased reliance on state support and prudent healthcare

A range of factors in western societies have medicalised the aging process, including the drive by pharmaceutical companies to sell more products, an increasing willingness to spend disposable income on health and a degree of denial of the consequences of aging\textsuperscript{29}. The Welsh Government has tried to counter this trend in its recent emphasis on ‘prudent healthcare’ and ‘coproduction’ but a sustained change of direction will require a sustained drive to achieve cultural change. Given the probability of ongoing financial stringency in the public sector this emphasis is likely to continue to be important. As the cost of caring for the elderly population grows, the state may not be able to supply all the needs of those who require support and a greater proportion may have to be provided by families, communities and the voluntary sector\textsuperscript{30,31}.

The high proportion of deaths occurring in hospital also places a demand on secondary care services as a whole and appears to be related to the societal changes in relation to death and dying\textsuperscript{32,33}. Wales has developed guidance around some aspects of this issue including addressing inappropriate attempts at cardio-pulmonary resuscitation (DNR-CPR), but it could be argued that there is still over-treatment of some patients with a poor prognosis and a failure to switch from curative to palliative services in a proportion of patients who die in hospital\textsuperscript{34}. Unscheduled care admissions of patients who would be better provided with hospice care or high quality palliative care at home is a major systems level challenge that needs to be urgently addressed to free up resources for other unscheduled care work.


4.3 Structural complexity

The flow of patients into unscheduled care services is complex. However, the system as a whole can adequately be monitored by an analysis confined to just the main flows. The Unscheduled Care Dashboard has therefore restricted analysis to In-Hours and Out-of-Hours primary care, Ambulance Services, NHS Direct, and Emergency Departments. There is a case for capturing some additional flows, particularly around community care and social care. A full mapping of patient flows was developed for the Unscheduled Care Board in Wales in 2011 and is shown in figure 8.

**Figure 8: Model of patient flow into unscheduled care in Wales**


The public, and sometimes even professionals, are not fully aware of all these pathways in the unscheduled care pathway and how best to interact with this complex system. This contributes towards inappropriate demand at the wrong point in the care pathway, for example, phoning NHS Direct when an ambulance is clearly needed adds an extra step and utilises unscheduled care resource. This issue was raised in relation to the
creation of NHS Direct but has never been adequately addressed\textsuperscript{35}. Another example associated with the increasing complexity of the system is the shift to telephone based care. There is some evidence that contrary to what one might expect, increased use of telephone care may result in an increase in the use of unscheduled care resources compared to face-to-face care\textsuperscript{36}.

4.4 System rigidity

System rigidity impacts on the management of fluctuations in unscheduled care demand and is one of the major causes of the current gap between supply and demand. The increasing codification of healthcare processes and procedures over a number of decades has made the system increasingly rigid\textsuperscript{37}, for example, rules that compel staff to take all their holidays within the financial year or else lose their entitlement to it, increase staff leave towards the end of the financial year reducing flexibility in service planning. Another example of increasing system rigidity includes aspects of the 2003 Consultant Contract. Greater central managerial control was obtained, but perhaps traded off against flexibility and local responsiveness, which it can be argued has had an adverse effect on the overall responsiveness of the system to surges in demand for unscheduled care.

As is pointed out elsewhere, there is a trade off between efficiency and resilience health systems which is particularly relevant to the planning of staffing levels and work patterns\textsuperscript{38}. Planning staffing levels on the basis of ‘average demand’ does not provide the spare capacity required to meet ‘peak demand’. Statistically speaking, planning on the basis of ‘average demand’ will result in staff that are over stretched half the time\textsuperscript{39}.

Although the evidence is mixed, there is some evidence that staff become tired and work less efficiently towards the end of twelve hour shifts as

\begin{itemize}
\item \textsuperscript{35} George, S. (2002). NHS Direct audited. BMJ, 324(7337), 558-559.
\item \textsuperscript{37} Ranji SR. (2014). What Gets Measured Gets (Micro) managed. JAMA 312(16), 1637-1638.
\item \textsuperscript{38} Hutchinson A. (2014). Exploring safety, quality and resilience in health care (Doctoral dissertation) Sheffield: University of Sheffield.
\item \textsuperscript{39} Andrew J. (2014) trusted to care Report of the external independent Review of the Princess of Wales Hospital and Neath Port Talbot Hospital at Abertawe Bro Morgannwg University Health Board. Cardiff: Welsh Assembly Government.
\end{itemize}
compared to eight hour shifts\textsuperscript{40}. However, health services are generally moving towards twelve hour shifts as the norm for nursing staff.

It has been argued that the increased codification of clinical contractual obligations has reduced trust between clinical staff and NHS management\textsuperscript{41}, which has resulted in front line staff being less willing to “go the extra mile” to make the system work.

\textbf{4.5 Inadequate investment in primary and community care as opposed to secondary care}

The growth in hospital consultants has been around 120\% between 1996 and 2012. There has also been a growth in management of around 50\% over the same time period. Over the same time period the growth in the GP head count has been 14\%. However, this appears to have been accompanied by a growth in part time work in General Practice, which would mean that the growth in primary care is less than the headline 14\%. The relative changes are shown in figure 9.


Figure 9: Percentage change in NHS staff groups General Practitioners and Hospital Consultants in Wales, beds and population served 1989 to 2012 (1996 base)

Source: Generated by Dr Mark Temple from data provided by Health Statistics Wales, Welsh Government.

There is evidence from NHS staff in different parts of Wales that the number of GPs on call at any one time providing Out-of-Hours care has fallen steadily since the change in the GP Contract in 2003. This appears to have generated a mismatch between supply and demand which has become a structural problem across the healthcare system and has pushed care towards Emergency Departments.

Other changes in the structure of primary care are also relevant. GPs practices were strongly encouraged to move to appointments-based systems over the last three decades. This primarily benefits those who are generally fit and healthy and in work rather than the young, the elderly, or those with long term conditions. This structural change reduced the same-day surge capacity in relation to unscheduled care demand compared to the historic model of open surgeries without appointments where patients could turn up, wait and be seen the same day, which prioritised unscheduled care demand.
4.6 Continuity and discontinuity of care

There is evidence that continuity of care is associated with reduced emergency admissions\(^{42}\) and planned admissions and significant cost savings\(^{43}\). Yet over the last forty years there has been less and less continuity of care in primary, community and secondary care sectors.

Continuity of care in terms of seeing the same GP at each visit has been eroded over a number of decades by system redesign. Evidence suggests that providing 1% of patients greater access to their ‘own doctor’ is associated with an annual cost saving of £20,000 per year in the health service and better health care quality\(^{44,45}\). A recent RCGP report estimates that significant savings could be made by diverting patients from Emergency Departments to Primary Care\(^{46}\). There is clearly a case for the incentivisation of greater continuity of care in primary care\(^{47}\).

There has been a shift in the delivery of primary care from single-handed GP practices, (where doctors were very well acquainted with the history of each patient and who had long-established relationships of trust with patients) to large group practices where objective standards of care are better but inter-personal relationships are poorer. The same has happened in district nursing, which has moved to larger teams and less of a link between specific nurses, local GP practices and a set of well known patients. In secondary care, recent anecdotal evidence indicates that in some hospitals, when a patient moves to a different ward because of bed shortages, the care of the patient moves to a different set of junior doctors and a different consultant. A significant proportion of complex elderly patients in some hospitals appear to move from one ward to another several times during one admission. The result is the loss of a sense of “ownership” of the problems of a patient by specific medical and

---


nursing staff, loss of continuity of care and loss of trust this leads to long lengths of stay and poorer outcomes.

The reduction in new-to-follow-up ratios in hospital outpatients has increased efficiency. However, it has reduced continuity of care. Historically, secondary care consultants in district general hospitals had a pool of more severe patients with relevant conditions that they followed up on a long term basis. This gave a sense of long term shared care with GPs for these patients. Problems could be resolved by a phone call because both consultant and GP were familiar with the patient. The use of Community Geriatricians may help to address this issue⁴⁸.

This issue is picked up in the Francis⁴⁹ and Andrews⁵⁰ reports (i.e. a culture where staff set up a mental boundary so that the needs of some patients are seen as being outside their responsibility). This is possibly a symptom of an over-stretched health care system.

### 4.7 The efficiency/fragility paradox

The Western world has had a sustained drive for efficiency across all sectors, including health, with annual efficiency targets and a philosophy of Just-in-Time, Lean Thinking and reduction in waste and variation. This work has had admirable benefits, particularly in relation to planned care. However, what appears to have been lost sight of is that there is an inverse relationship between efficiency and resilience⁵¹. In others words highly efficient pathways where variation has successfully been reduced, are potentially more fragile if they cannot flex to address sudden variations in demand⁵².

The fragility underlying efficiency was demonstrated a few years ago when the UK had a petrol strike and it materialised that the UK only had reserves to last around three days. The same applies to unscheduled care. There are a wide range of determinants of unscheduled care which

---


have a high intrinsic variance. Each of these determinants can amplify the effect of other determinants of unscheduled care, means that the variation in demand is sometimes highly amplified. Unusual levels of demand for unscheduled care occur more frequently than would be expected if the variation in unscheduled care was normally distributed. The only way to cope with such non-linear systems is to have significant levels of spare capacity, or ‘head room’ built into the system which can be brought into use very rapidly. Examples of systems that require spare capacity include the fire service and the electricity supplied by the National Grid. Both systems need to be able to respond rapidly to sudden or unexpected surges in demand.

4.7.1 Over-specialisation of secondary care staff

Over a number of decades there has been over-specialisation of clinical staff in secondary care. As a result, despite a significant expansion in consultant numbers, there has been a fall in the proportion of doctors involved in unscheduled care. The Greenaway Report\(^53\) recognises this problem and states that, “All doctors must be able to manage acutely ill patients with multiple co-morbidities within their broad specialty area”. Health Boards could take actions to accelerate the process by which all secondary care consultants contribute towards meeting some aspect of unscheduled care demand.

4.7.2 Shift in focus form unscheduled care to planned care

It can be argued that the NHS has moved from its original focus on ‘unscheduled and emergency care’ to a focus on ‘planned care’ and that there has been a shift in the skill mix of those employed by the NHS, with greater emphasis being placed on the development of the skills required to address ‘planned care’ rather than ‘unscheduled care’.

4.7.3 Challenges related to ‘flow’ through diagnostic services

There are multiple bottle necks in the care pathway for patients. A significant proportion of these bottlenecks relate to diagnostic services. The 1,000 Lives flow programme is supporting work on these pathways and they have therefore not been addressed more fully in this document. One systems level issue that does arise is the differential access of primary and secondary care to certain diagnostic services. In some cases this drives unscheduled care demand to Emergency Departments as patients try to use this to ‘jump the queue’ and access a diagnostic test.

4.7.4 Falling effectiveness when close to capacity

The dramatic fall in effectiveness of a service when it is close to capacity is an important but often overlooked phenomenon. The effect can be seen on the motorway. When the traffic reaches a critical volume, the traffic temporarily comes to a standstill despite the fact that there has been no accident, and even when the traffic starts to move again it does so at a very slow pace. When hospitals reach capacity this “treacle-like effect” has profound effects on flow through diagnostic and treatment facilities. Work by Rod Jones provides information on safe occupancy levels for different sized units. Occupancy levels above 80% are generally associated with reduced effectiveness in general medical and surgical bed pools. Smaller units in the pathway such as emergency department bays, medical and surgical assessment units have an optimum occupancy of around 60%. The sequential nature of pinch points in the patient care pathway are also a factor leading to internal queues and problems with “flow”.

The probability of a queue developing, given the size of a unit is shown in table 3. The likelihood of a queue forming is more profound in smaller units.

Table 3: Probability of a queue developing for different sizes of bed pools

<table>
<thead>
<tr>
<th>Size of Bed Pool</th>
<th>Probability of Queue Developing</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.1% 1% 5% 20% 50%</td>
</tr>
<tr>
<td>30% occupancy</td>
<td>44% occupancy 59% occupancy 78% occupancy 92% occupancy</td>
</tr>
<tr>
<td>65% occupancy</td>
<td>76% occupancy 85% occupancy 94% occupancy 98% occupancy</td>
</tr>
<tr>
<td>74% occupancy</td>
<td>83% occupancy 91% occupancy 97% occupancy 99% occupancy</td>
</tr>
<tr>
<td>88% occupancy</td>
<td>92% occupancy 96% occupancy 98% occupancy 99% occupancy</td>
</tr>
</tbody>
</table>

Most hospital bed pools are small. For example, many Medical and Surgical Assessment Units have a bed pool of around 10-20 beds, with a consequent 20% probability of a queue for admission developing even if

the unit runs at an average occupancy of 80-85%. That is to say, we would expect a 10 bed assessment unit to be unable to immediately admit one in five patients. At 99% occupancy even a pool of 50 beds will be unable to immediately admit one in two patients. Hospital units in Wales are regularly running at 99% occupancy, and there can be little doubt that the subsequent delays in receiving care must be having adverse effects on outcomes in some patients. There is a serious case for considering more traditional admissions systems, where patients are admitted to the whole pool of hospital beds, as such a system would be more resilient.
4.7.5 Hospital beds in Wales

There has been a steady fall in the number of hospital beds in Wales over several decades as is shown in figure 10.

**Figure 10: Changes in bed numbers in Wales, 1989-90 to 2013-14**

The number of beds in Wales has fallen from 19,855 in 1989-90 to 11,241 in 2013-14. The number of hospital beds per 1,000 population for a range of OECD countries is shown in figure 11.

**Figure 11: Hospital beds per 1,000 population across OECD countries, 2012**

![Figure 11: Hospital beds per 1,000 population across OECD countries, 2012](image)


Wales has fewer hospital beds per 1,000 population than many other OECD countries, although the number of hospital beds per 1,000 population is higher than that in the UK as a whole. Given the high proportion of the population over 80 years, one can argue that the cut in hospital beds may not have been matched by an adequate rise in support systems in the community and that there is a short term deficit in hospital beds, particularly when the number of nursing home beds in the community is taken into account.

In summary, across all resources a resilient system requires significant spare capacity to deal with expected fluctuations in unscheduled care demand.
4.8  **System rigidity and application of simple solutions to complicated, complex or chaotic situations**

As has already been pointed out, the systematisation of processes can increase efficiency, and a ‘production line’ approach is very suited to simple problems. However, the introduction of extensive guidelines and policies has resulted in a rigid system, which is highly efficient when things are going well but is intrinsically fragile when under severe pressure. Examples may be useful in describing the phenomenon. Many nursing homes have been criticised for missing a hip fracture when a patient falls out of bed. The result has been to create a ‘rule’ such as "every patient who falls out of bed must be seen by a doctor". This has had significant adverse effects. Over stretched Out-of-Hours services are commonly refusing to visit such patients; to ‘cover their back’ nursing homes then phone 999 for an ambulance to take the patient to hospital. Such a confused elderly patient often ends up with an inappropriate admission to hospital and in some circumstances ends up dying in hospital. The root cause here is a risk-averse system and a society that demands that no mistakes are ever made, leading to rules that protect one party but ultimately disadvantage the patient. This defensive practice, where the responsibilities for problems are pushed onto others, rather than being ‘owned’ and addressed, appears to be becoming part of NHS culture. Another example may be useful. Rigid rules have been introduced in places which have stopped NHS staff prioritising care for other members of staff. There are reports of physiotherapists who could not get physiotherapy for their injury and could not get back to helping an overstretched physiotherapy waiting list because of a rule that all staff must join the waiting list.

The “Blame Culture” that was recognised by the Francis and Andrews Reports as an important factor is related to the concept of sticking to the rules or perceived priority (completing paper work to avoid criticism) even when a situation calls for an exception to the rules or reprioritisation to meet the needs of patients. This phenomenon is related to decision-making rigidity where obeying the rules can inadvertently be perceived as more important than doing what is best for the patient. This can be described as ‘micro-politicisation’ of the interactions between NHS staff, where there is a culture which believes that observing the rules is more

---


important than doing what is best for the patient\textsuperscript{57}. Such a context encourages excessive codification of processes as opposed to trust and co-production.

A risk-averse and audit-orientated society clearly contributes to this phenomenon\textsuperscript{58,59}. Society seems to expect that there should be a rule to address every given situation and that uncertainty of outcome can be reduced by the imposition of process. There may be a link to the 24-hour nature of modern news media, which needs to be fed by a constant series of stories and a predilection for ‘health stories’ as these appear to boost public interest and generate revenue for media outlets.

The work of David Snowden is useful in this regard\textsuperscript{60}, as it recognises that simple solutions only work for simple problems and that for complex problems a much better paradigm is a series of iterative ‘best guesses’ followed by reassessment and adjustment of the initial approach. This used to be the primary historic approach to medical problems.

55 Service users’ beliefs

Research on the reasons for choosing to visit Emergency Departments (rather than use alternative NHS services) has been undertaken in North Wales\textsuperscript{61,62}. This research has identified some push factors, related to perceived issues with General Practice, Out-of-Hours services, minor injuries units, NHS Direct and pharmacies and also some pull factors which made Emergency Departments attractive. Similar factors have been identified by other researchers\textsuperscript{63,64} These are summarised below.

\begin{itemize}
\item \textsuperscript{61} Atenstaedt R et al. (2014) Exploring the reasons why patients attend the Emergency Department for non-urgent treatment. Bangor: Public Health Wales.
\item \textsuperscript{62} Social Change UK. (2014) Choose Well: Attitudes, values and behaviours of adults in Rhyl, Caernarfon and Wrexham in relation to choosing the most appropriate NHS services for treatment.
\end{itemize}
Perceived issues with Primary Care

- Inconvenient hours
- Inconvenient location (e.g. not near workplace)
- Cannot get an appointment when wanted (same day)
- Perception that GPs not knowledgeable / refer to hospital anyway
- Never get to see same GP (who knows the family history)
- Variation between practices (different opening hours, appointments policy, etc)

Perceived issues with Out-of-Hours services

- Lack of awareness (where, how to access)
- Confusion with NHS Direct

Perceived issues with Minor Injuries Units

- Location not always convenient (further away than ED)
- May get sent to ED anyway so waste of time / additional travel
- Lack of awareness (location, opening hours, services, when to use)
- X-ray not always available

Perceived issues with NHS Direct

- Cost of calling (especially from a mobile or if held in a queue)
- Long waits (on-hold or waiting for call return)
- Lack of awareness (that service exists, what it’s for, the number to call)
- Poor experience in past / perception based on others’ experiences

Perceived issues with pharmacies

- Pharmacies are rarely referred by patients as a place to go for help and support despite them being visible, accessible and available in the heart of many communities
- There appears to be low awareness of how a pharmacist can help

Some of the factors drawing individuals to Emergency Departments

- Lack of information, low awareness or poor understanding of alternatives to ED
- Multiple NHS re-organisations that have lead to confusion of availability & relevance of alternative services
- Passing on of responsibility by different providers
- The need for reassurance (especially for children) and greater confidence in hospitals (experts and equipment which would give peace of mind)

---

• A perceived lack of access to the patient’s records at other services
• Influence and advice from family, friends and social media
• The high profile of Emergency Departments and the universal awareness of their location, services, and 24 hour accessibility
• A possible need for an x-ray which could not be obtained at a GP surgery or some Minor Injuries Units

6 Conclusion

This report has identified a range of environmental, behavioural and structural factors that contribute towards rising demand for unscheduled care in Wales and the gap between supply and demand. The unscheduled care system in Wales is complex, with each driver interacting with other drivers in the system, potentially amplifying the effect of other drivers.

It is difficult to determine the relative proportion of avoidable demand provided by each determinant that has been referred to. In terms of short term demand drivers the biggest challenge is a mismatch between the resources deployed at any given time and the demand at that time. There are also issues of systems efficiency and service quality. These factors may be addressed by work on ‘flow’ using healthcare improvement methodologies. Weather patterns and seasonal effects can cause significant disruption if systems are run at close to capacity. Short term flexing of capacity may be the most effective way of creating a more resilient system with the ‘headroom’ to manage this issue of occupancy rates.

The biggest driver of demand in the long term is failure demand, that is, “demand caused by a failure to do something or do something right for the customer”. The current system inadvertently pushes patients from primary care and pulls them towards secondary care. A toolbox of interventions to address some of these factors has been developed by Public Health Wales to support Health Boards.

Patient behaviours in relation to inappropriate Emergency Department use are also important and can be partially modified by educational


approaches such as “Choose Well”, but are probably best addressed in the longer term by making other options more attractive to patients.

It is important not to lose sight of societal inequalities as an underlying driver of demand, although the response to this issue requires much broader societal engagement.

7 Recommendations

The primary purpose of this report is to analyse demand drivers rather than provide recommendations. However, three priority areas where there are modifiable demand drivers were be identified.

1. Operating systems with adequate ‘head room’, either by operating at lower occupancy rates or by having rapid surge capacity of up to 50% in terms of beds and staff in units that are key points along the patient pathway.

2. A shift to ‘pull through’ rather than ‘push through’ systems at each stage in a patient care pathway. This would mean that social services and reablement teams were proactively identifying patients to take out of hospital and back into the community, and medical wards were proactively identifying patients who could be taken out of Emergency Departments, Intensive Care Units and assessment Units.

3. Rebalancing the relative priority of ‘unscheduled’ and ‘planned care’, particularly in relation to the up-stream resource in primary, community and social care that can rapidly respond to fluctuations in demand.

A separate report “A Toolbox of Actions to Address Pressures in Unscheduled Care” provides a fuller list of evidence based interventions designed to address unscheduled care. “An Atlas of Variation in Healthcare” and a number of other reports have also been provided by Public Health Wales to support analysis of unscheduled care need in different areas across Wales and are available on the Public Health Wales website.