The Demography Myth: How Demographic Forecasting Underestimates Hospital Admissions, and Creates the Illusion that Fewer Hospital Beds and Community-based bed Equivalents, will be Required in the Future

Neeraj Beeknoo\textsuperscript{1*} and Rodney P. Jones\textsuperscript{2}

\textsuperscript{1}Kings College University Hospital, London, UK.
\textsuperscript{2}Healthcare Analysis and Forecasting, Worcester, UK.

Authors’ contributions

This work was carried out in collaboration between both authors. Author NB proposed the study. Author RPJ conducted the literature search, sourced the data on NHS beds and deaths in England, and prepared the first draft of the manuscript. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJMMR/2017/29984

Editor(s):
(1) Angelo Giardino, Texas Children’s Hospital, Houston, Texas and Pediatrics, Baylor College of Medicine, Houston, TX, USA.

Reviewers:
(1) Ooi Chee Kheong, Tan Tock Seng Hospital, Singapore.
(2) K. V. Ramanath, Rajiv Gandhi University of Health Sciences, India.
(3) Robert B. Tate, University of Manitoba, Canada.
Complete Peer review History: http://www.sciencedomain.org/review-history/17161

Received 8th October 2016
Accepted 28th November 2016
Published 8th December 2016

ABSTRACT

\textbf{Aims:} This Policy Article, which includes data synthesis, seeks to evaluate the effectiveness of demographic-based forecasting of future hospital admissions, and hence for hospital bed numbers. A role for the absolute number of deaths, as a proxy for persons approaching the end-of-life, will also be investigated, especially as this is the principle factor behind the volatility in bed occupancy.

\textbf{Study Design:} Literature review plus supporting analysis of relevant trends.

\textbf{Place and Duration of Study:} Studies from a variety of countries, analysis of data relating to the NHS in England, additional analysis relating to the Kings College Hospital, London.

\textbf{Findings:} Demographic forecasting is subject to the constant risk fallacy, namely, admission rates are changing over time. A variety of factors can be seen to lead to changes in admission rates.

*Corresponding author: E-mail: neerajbeeknoo@hotmail.com;
Co-author: E-mail: hcaf_rod@yahoo.co.uk;
There is strong evidence that it is the trend in the absolute number of deaths, rather than demography which drives most of the marginal changes in bed demand. Up to 55% of a person’s lifetime bed utilization may occur in the last year of life. However, the ratio of deaths per 1,000 population ranges from 4 to 16 (inner city communities to more rural and retirement locations). Deaths per GP also ranges between 5 to 27, while the crude mortality rate in English hospitals ranges from 1.3% to 5.4%. These imply that in some locations end-of-life has a greater impact not only on primary, secondary and social care resources, but also on the volatility in resource requirements. There are 20 deaths per GP in the Torbay integrated health and social care organization indicating that the 30% fewer bed days per death achieved in this organization may not be replicated elsewhere. Same-day-stay admissions require separate forecasts since they rely on technological trends driving day surgery rates (elective admissions) and the need for observation and rapid diagnosis in medicine (both elective but mainly emergency admissions). In England, an obsession with fewer beds on behalf of the Department of Health (now NHS England) has been a key factor bringing the NHS into crisis. The worst affected hospitals now run at close to 100% average occupancy. However, the period 22 weeks prior to death appears to mark a watershed for acute care, where surgical, medical and critical care interventions become increasingly futile. A shift to a more palliative-based model of non-acute care is recommended, and will lead to considerable savings in occupied beds. A new generation of models are desperately required, as are supporting biochemistry, vital sign and frailty-based algorithms to enable care to be diverted out of acute hospitals during the last weeks of life.

Keywords: Hospital planning; activity forecasting; hospital bed numbers; demographic forecasting; age standardization; palliative care.

1. INTRODUCTION

This discussion paper arises out 25-year careers associated with the English National Health Service (NHS). Many of the issues will be applicable in other countries, however, it is unavoidable that parts will be colored by issues specific to the English NHS. It is therefore useful to provide a brief overview of features peculiar to the English NHS.

The UK NHS is run as a devolved service across four separate countries (England, Northern Ireland, Scotland and Wales). Each country of the UK can vary the funding directed toward their part of the NHS, and will have different management styles and a different emphasis on the balance between centralization/decentralization. Common to all is a tight rein on capital expenditure, and a centrally run approval process for large capital projects which are almost exclusively (except in more recent times for Scotland) provided by the Private Finance Initiative (PFI). Revenue expenditure is also tightly controlled and NHS organizations are expected to achieve break-even. In England, those who fail to achieve break-even will eventually go into ‘special measures’ or be merged with nearby organizations. Hence many of the insights in this discussion piece emanate out of a health care system attempting to grapple with very tight control of capital and revenue costs, which is a common theme around the world, but will have ‘local’ implications for the NHS in England. As in other countries there will also be examples of the unintended consequences of government policy.

Hence, if you were to ask any hospital manager in England regarding how to forecast admissions, they would nearly all reply that the method of choice is demographic-based using current admission rates per age band multiplied by future population changes. Huge financial spreadsheets are constructed in almost every NHS hospital based on demographic forecasting, and the assumption of demographic-based trends is standard in the software supplied to the NHS by outside companies. Hospital bed planning is contingent on demographic activity forecasting, and future hospital bed requirements then apply an estimate of future length of stay, shifts from overnight elective to day surgery, and an occupancy margin. The same observation would probably apply to many countries around the world.

While few managers within the NHS seem to be aware that this method may have serious limitations, and may even grossly underestimate future demand, a survey of UK academics would probably yield a far wider appreciation of the multiple factors contingent in the trends in
demand. This discussion piece therefore attempts to walk the tightrope between wider academic issues and a pragmatic discussion directed at management and policy makers.

For example, Fig. 1 shows the trend in average midnight occupancy for acute hospitals in England. As can be seen the last time the NHS in England averaged 85% occupancy (often considered to be the safe average occupancy margin [1]), was during the summer of 2011. In the quarter ending June 2016, some 17% of acute hospitals were running at or above 95% average midnight occupancy [2]. The term ‘midnight occupancy’ is emphasized simply because midnight is near to the point of minimum occupancy in the 24-hour occupancy cycle, and close to 100% occupancy is common across the whole of England during week-day day-time hours, even during the summer months. The situation has worsened over time, with average acute midnight occupancy of 83.1% back in the 2000/01 financial year [2]. Overnight bed numbers have declined by 24% since 2000/01, and this decline has partly contributed to the rising occupancy levels. Clearly the forecasting of demand and bed requirements has gone seriously astray.

This discussion piece will therefore investigate the limitations of the demographic methodology and make pragmatic suggestions regarding the forecasting of future admissions and bed requirements. Many of these issues have already been raised in discussion of NHS bed planning methodologies published in the early 2000’s [3,4], however this policy piece will seek to cover developments since that time. Indeed, the emphasis is to explore issues which are not commonly addressed in the wider literature, and to highlight how death acts as a direct estimate of end-of-life use of acute services – and how this fact may be used to eventually reduce the need for acute hospital care.

2. AGE STANDARDIZATION

Age standardization was introduced in 1844 to compare the death rates in various ‘sanitary districts’ with markedly different age structures [reviewed in 5]. No one would deny that this is an entirely appropriate application of the methodology.

![Fig. 1. Average quarterly midnight occupancy for all acute NHS hospitals in England](https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/bed-data-overnight/)

At some point in hospital planning someone must have translated the principles of age standardization into demographic-based forecasting of hospital admissions. Namely, take today's admission rate in each age band, along with a forecast of future population in each age band, sum by age band to give future admissions, multiply by future expected length of stay (LOS), and apply an occupancy margin to give required beds [3,4]. As an example, a short article in the British Medical Journal in March 1981 discusses the robustness of a health plan for London involving a 23% reduction in hospital bed numbers [6]. On that occasion the discussion centered around the uncertainty in population forecasts by age band, however similar unreliability in other assumptions was likewise alluded to [6].

However, unlike age standardization, there is no scientifically credible basis for this approach, as the sole descriptive factor in forecasting admissions.

3. DEMOGRAPHY FAILS TO EXPLAIN MOST TRENDS

As was elegantly pointed out in the seminal paper of Nicholl [7], the rates per age band are subject to the 'constant risk fallacy', i.e. that they are not constant but change over time. This observation has been repeatedly confirmed for admission rates. Hence a Scottish study between 1981 and 2001 showed that rates of emergency admission in all age bands had risen in a near-linear fashion over time, with the rate of increase increasing with age [8]. A study conducted in England between 1989/90 and 1995/96 concluded that only a small proportion of the increase in emergency admissions was due to demographic change [9]. Another English study conducted between 2004 and 2009 demonstrated that only 40% of the increase in emergency admissions could be explained by changing population demography [10]. A study of English admission rates (emergency + elective) among those aged 75+ between 1998 and 2011 showed that growth was highest in a group of 90 high volume medical diagnoses typically associated with multi-morbidity clusters, i.e. persons with condition A, will also have a higher risk of having conditions B, C and D [11]. In Germany between 2000 and 2008, changing disease incidence, hospitalization rates and patterns of age all interact to give the observed trend in the total number of admissions for different chronic diseases [12]. Deeper analysis reveals that admissions (elective + emergency) in the surgical group of specialties does indeed roughly follow a demographic trajectory [13]. Trauma admissions are rising at a rate slightly lower than demography, but with long-term cycles seemingly associated with metrological variables [14-16]. Elective orthopedics is rising at a rate moderately higher than demography, due to ongoing expansion in the sophistication and range of orthopedic procedures, and the need for revision of previous joint replacement surgery [13], however, the major part of the unexplained rise is confined to the medical group of specialties [13].

3.1 Non-demographic Factors

There are a number of non-demographic factors which may affect admission rates and length of stay such as, more investigations and treatments of increasing complexity are now available, more co-morbidities are now present and also more often recognized and addressed, more fragmentation of expertise and care with attendant communication issues, increasing expectations of patients and their relatives, with more rapid and demanding seeking of treatment, increasing risk aversion of clinicians lowering the threshold for admission, more iatrogenic illness, organizational and structural issues, i.e. too much centralization, too much dependence on trainees, poor training and supervision, and reducing mortality has increased morbidity and need for institutional care.

No one is seeking to deny that these issues are relevant in modern medicine, and are contributing to a background level of increase in the non-demographic component of growth. Indeed, these form part of the conundrum of modern medicine and will require their own solutions. However, the emphasis of this discussion piece is that up to the present there are very few studies seeking to disentangle to what extent end-of-life contributes to admission rates or length of stay rather than age per se, or if many of the above only have relevance in say the last year of life. It is hoped that this discussion piece will make some contribution to disentangling these issues.

4. THE ROLE OF AGE

It is well known that in all Western Health Care systems that the number of older people is
increasing due to increasing longevity. It is therefore common to hear the statement that health services are under pressure due to ‘the ageing population’. Fig. 2 shows how age does indeed contribute to the total number of admissions. The spike at age 0-4 is largely due to neonates and children in their first year of life, while slight bulge between 20 and 39 is due to birth related admissions. Tables A1 and A2 in the Appendix list the top 25 reasons for admission in the oldest and youngest groups during 2012/13. However, the trend is dominated by an increasing admission rate as age increases. So in one sense, the ageing population will contribute to higher total admissions – however, if demographic forecasting was working correctly any increase should have been anticipated and planned for.

However, a study in Sweden for persons over the age of 60 has demonstrated that the age at first admission increased in line with rising life expectancy [17], hence age per se may be slightly misleading, and this issue will be discussed later.

At this point we need to address the issues behind the seemingly small differences between the three years in Fig. 1.

5. MORE ENIGMATIC TRENDS

While a study of long-term trends is useful in establishing general principles, it is sometimes useful to investigate the minutiae to see if more enigmatic trends may lie hidden from view. In this respect Fig. 3 examines the seemingly trivial differences in admission rates between the three years seen in Fig. 2 – under the assumption that something which is not understood will often be ignored as an artefact.

Fig. 3 clearly shows that age per se may not be the primary factor, as has been assumed to date. There appear to be systematic differences due to age which appear to defy simple explanation. The numbers involved are simply too large for the differences to arise from Poisson variation. Indeed, the question needs to be asked if the use of standard 5-year age bands is masking further complexity?

**Fig. 2. Consultant episodes (elective + emergency) per 1000 head of population in England by age band**

*Data for consultant episodes by age band is from Hospital Episode Statistics (HES) data available from the NHS Digital website. Mid-year population estimates by age band for England is from the Office for National Statistics website and is on a calendar year basis. Data includes mental health (less than 5 per 1000 across all age bands) and maternity (74 per 1000 at age 29-34, less for other ages)*
6. DEATHS AND ADMISSIONS

Discussion up to this point has centered around the role of demographic change, however, it has been repeatedly noted that lifetime utilization of hospital services is usually concentrated in the last year of life, irrespective of the age at death [18-24]. While this appears to contradict any demographic basis for forecasting admissions, many elective procedures (which account for around 70% of admissions in England) are not immediately life-saving, and occur in generally younger patients. Also, because of the effect of age on the likelihood of death (see Fig. A1 in the Appendix), the end-of-life contribution can be roughly approximated by the age band approach employed in Fig. 2. See further discussion in the section dealing with admissions per death.

However, Fig. 4 demonstrates how the point at which 50% of deaths have occurred has shifted over time. Put another way the average 81-year-old in 2015 is just as healthy as the average 72-year-old back in 1963. Demographic forecasting using 1963 age-related admissions would be totally ineffective for predicting 2016 admissions, i.e. the outworking of the constant risk fallacy.
Fig. 4. Age at 50% of deaths in England and Wales

Age at 50% of deaths was determined using single year of age death statistics for England & Wales derived from the Office for National Statistics.

A Scottish study demonstrated that the fact of a hospital admission demonstrated a high risk of death within the next year, with 46% of persons in hospital at the 31st March 2010 (surgical or medical) aged 85+ being dead within one year. A medical admission had double the risk of death within one year [28].

The principle reason that death acts as a good proxy for end-of-life is the observation that admissions begin to rapidly increase in the last 22 weeks of life, and especially escalate in the last eight weeks of life [24]. These observations have been replicated using Dutch health insurance claims [29]. Hence while the exact time-profile up to death may vary for different conditions the 22 week break point is a good population average.

However, the key point is that it is the total number of deaths (rather than the more usual age-standardized death rate), that is an excellent proxy for the number of individuals in the high-admission, end-of-life part of the life cycle.

Having established the role of deaths as a proxy for end-of-life healthcare utilization, it is useful to examine if deaths show the same patterns observed in Fig. 3.

6.1 Single-year-of-age Effects

Fig. 5 therefore presents a similar analysis to Fig. 3 for deaths in England relative to the 2012 calendar year which has been chosen as the closest approximation to the 2012/13 financial year data available for admissions. However, on this occasion single-year-of-age is used to illustrate some important points.

There are two potential artefacts in Fig. 5. These involve persons born during the World War I and II baby booms who are aged 66 and 92 respectively in 2013, with more survivors than for those born one year earlier. Hence in 1920 compared to 1919 there were 38% more births, while in 1946 there were 21% more births, with knock-on effects to survivors in 2012 onward. The trend in Fig. 4 has therefore been standardized against single year population in each year. The point of interest is that a saw-tooth pattern emerges with an approximate 4-year interval from trough to trough. This saw tooth pattern is even reflected in emergency department attendances [30]. This same saw-tooth pattern is also observed when single-year-of-age is applied to unexpected changes in medical emergency admissions [31-35].

A further point of interest in Fig. 5 is an influenza outbreak in early 2015 largely restores deaths to 2013 levels for those aged 70+, but with sometimes dramatic single year differences below this age. The significance of this will be explored in the next section.
6.2 Antigenic Original Sin

It is of interest to note that such saw-tooth patterns arise in the phenomena called ‘antigenic original sin’ [36-38], whereby the immune system is primed by past exposure to pathogens having multiple strains and variants. The immune response to successive exposure is heavily primed to the first strain encountered, usually during childhood. This immune priming can be helpful (provided there is significant cross-immunity) or unhelpful. Provided the person survives the infection with subsequent strains, the immune system will eventually mount an appropriate immune response leading to a complex array of antibody landscapes, which are subject to time decay [39]. Hence the saw-tooth patterns which depend on age at first and subsequent exposure.

How does antigenic original sin influence deaths and admissions? A good example is the Swine flu (H1N1) 2009 epidemic. Since influenza virus A (H1N1) circulated continuously between 1918 and 1957, most people born before 1957 had been infected with a H1N1 subtype. Due to relatively good cross-immunity between the strains persons aged 52 to 91 therefore largely avoided the effects of the 2009 Swine flu epidemic, and effects were therefore greatest in those aged under 52 [38]. Fig. A3 in the Appendix shows an estimate of the additional medical and pediatric admissions in England which occurred in the 2009/10 Swine flu year (after adjusting for underlying growth), where it can be seen that admissions in the under 50’s showed the greatest increase.

The effect of the January 2015 influenza outbreak relative to 2010 in Fig. 4 can be interpreted in the same light. However, the overall effect of the 2015 influenza event is more complex since the population was immunized with an incorrect antigen mix, and that the influenza outbreak appeared to interact with an earlier outbreak of another (unidentified) agent – hence possible original antigenic sin effects against two pathogens, complicated by inappropriate vaccination against influenza [discussed in 40-41].

6.3 Roles for the Birth Rate

While the number of births each year has an obvious role relating to maternity, neonatal and
pediatric services, there are less obvious roles relating to the acquisition of immunity during childhood, and the consequent transmission of infections to adults and the elderly [42-43]. In this respect Fig. 6 presents the trend in number of births in England and Wales from 1938 to 2015. Note the peaks arising from the World War I and II baby booms, and the subsequent mini-booms as the babies subsequently reach child bearing age. The mini-boom of the 1980’s was enhanced by inward immigration from the Commonwealth countries, while that commencing 2002 was enhanced by inward immigration from the European Union Accession Eight countries.

With over 1,400 known human pathogens [44], it is now recognized that many diseases, including cancers, and even obesity can have an infectious basis [45-50], and that other conditions can be exacerbated by unrelated infections. For example, secondary bacterial infection following a viral infection can lead to pneumonia, while asthma is most commonly exacerbated by respiratory infection, especially mixed viral and bacterial infections [51,52].

The large waves in the number of children can then lead to corresponding waves in the transmission of a variety of infections, which may lead to changes in the incidence of various conditions in the ensuing years. Such waves are called cohort effects, and can also arise from non-infectious sources such as smoking prevalence, etc.

It is of interest that cohort effects are known to influence the death rates in particular years [53-54], and by extension the rates of emergency admissions associated with end of life. It is also well known that the conditions related to cause of death change over time [55]. In this respect, some conditions lead to higher bed utilization, and an Australian study established the following ratios of bed-days per death for various lifestyle choices: tobacco (43), illicit drugs (83), all causes (134), other causes (149), alcohol (200) [56].

Clearly, simple demographic forecasting is not able to encapsulate such complexity in the trends, especially in the end-of-life related component.

6.4 Using 2012 as a Base Year

The choice of 2012 or 2012/13 in Figs. 3 and 5 as the base year is not random, since both deaths and medical admissions experienced a totally unexpected and large increase in this year [29-33,57-61]. Those suffering from neurological disorders were especially prone to dying [59-61], and whatever caused this event spread across the entire UK in a manner expected from a relatively difficult to transmit persistent infectious agent [31-35,62-63]. Hence if the suspected infectious agent exists as a series of strains or variants then 2012 will set a precedent for subsequent years, especially 2014 or 2014/15 when another outbreak of the agent appears to have occurred [64-65].
6.5 Implications to the Use of 5-year Age Bands

Clearly something highly unusual is happening, which may have an infectious origin. The key point relating to demographic-based forecasting is the use of 5-year age bands. In this respect, 5-year age bands are an international standard in age standardization (which appear to have emerged in the 1930’s, perhaps as a compromise relating to the ability to accurately forecast the base population), although 10-year bands are sometimes used and the first age band is often broken down to age 0 and 1-4. However, if age bands are used in a context where antigenic original sin (or a similar phenomenon) is active in the population, it is highly likely that valuable information regarding fundamental mechanisms may be overlooked, and admission rates per age band will be subject to skewing. This is not a new concept, and a paper in 1983 by Holford [66] discussed the role of age bands, periodicity and cohort effects on the calculation of rates for vital events such as deaths.

In conclusion, both admissions and deaths show evidence for saw-tooth patterns when compared against 2012 as a base year – although this effect is not limited to 2012 [29,62,65]. Such saw-tooth effects may act as confounders when 5-year age bands are used to establish a base year for future forecasts. Extreme caution is therefore required in exactly quantifying any demographic-based component of growth.

However, such matters aside, it appears highly likely that deaths may play a far greater role in both the observed volatility, long-term trends in admissions, and occupied beds than has hitherto been appreciated. This issue will now be discussed.

7. DEATHS AND THE VOLATILITY IN ADMISSIONS AND BED DEMAND

A key observation is that demographic-based forecasting only ever generates smooth line trends. Even if demographic forecasts are augmented with all manner of societal and technological changes, these will still produce roughly smooth line trends. Such smooth line forecasts are entirely at variance with the known real-life volatility in admissions and bed occupancy [67,68], some of which was illustrated in Figs. 1 and 3.

To answer the question if absolute number of deaths (all-cause mortality), as a proxy for the number of persons in the last year of life, is reflected in the volume of hospital admissions (as measured using occupied beds), Fig. 7 therefore matches the number of acute occupied beds in England in each quarter with the number of deaths (all-cause mortality) occurring in each quarter in the interval 2010 to 2016. As can be seen there is an excellent match between deaths and the marginal change in admissions and hence occupied beds. Between 2012 and 2016, 100 deaths roughly match with 66 to 85 occupied beds. There is a degree of seasonality with highest number of occupied beds per death in the summer months. See Fig. A2 in the Appendix.

The trends revealed in Fig. 7 should be no surprise in view of the possibility that the last year of life may account for up to 55% of a person’s lifetime utilization of hospital bed-days. Clearly annual demographic forecasts are unable to reflect both the seasonal and end-of-life related component of bed occupancy.

7.1 Location-specific Volatility

Analysis of long-term trends in both deaths, admissions, costs and bed occupancy shows that some locations are subject to lower volatility than others [69-72]. This may be due to local weather patterns and other metrological variables, as has been observed in trauma admissions [16].

This raises the interesting question as to how individual hospitals may respond to the different profiles of death seen across the area from which their patients originate. This is illustrated in Fig. 8 for the Kings College Hospital (KCH) in London which receives the majority of emergency admissions from seven nearby Local Authority areas. A running 12-month total of deaths has been used to remove the month-of-year pattern in deaths (as in Fig. 5) and the expected trend for KCH has been derived based on the proportion of emergency admissions arriving from each Local Authority.

As can be seen the trend at KCH is relatively smooth compared to that of the individual Local Authority areas. Many of the peaks and troughs in the running 12-month totals at Local Authority level appear to cancel each other out, leaving a greatly moderated response to events in 2002, 2008 and 2014, and almost negligible response
to events in 2004, 2006, 2010 and 2012. This behavior is also reflected in the trends in emergency admissions at KCH, and appear to explain why other hospitals throughout the country exhibit far greater volatility in emergency admissions and bed occupancy over time.

Indeed, the response observed in London is less than that observed elsewhere in the UK, seemingly due to a generally younger population in London. These events have been proposed to be of an infectious origin [50], but whatever the cause they act to illustrate an important concept regarding the contribution from end-of-life.

Fig. A4 in the Appendix presents an interesting comparison of the disparity in the trends in deaths in the interval from the 12-month period ending December 2001 through to the 12-month period ending July 2016. As can be seen, inner London (which includes many of the Local Authorities in Fig. 6) shows a continuous trend downward, while at the other extreme Milton Keynes shows a trend upward. Milton Keynes was one of several new towns established in the late 1960’s, with continuous population growth since that time. Growth in the elderly population is escalating as residents (now in their 50’s and 60’s) relocate their older parents to Milton Keynes.

Inner London is the exact opposite to Milton Keynes, with one of the youngest populations in England. Persons who reach retirement age generally exit from London to retire in more peaceful locations such as the South West. The resulting ‘spare’ housing is then filled with an influx of younger people at the lower end of the housing ladder.

Clearly the end-of-life component of admissions, bed requirements and costs will be very different to that seen in London, and in theory, this should be reflected in the English Capitation Formula used to distribute funds throughout the NHS. Sadly, this is not the case, and to reemphasize it is the total number of deaths, and not the age-standardized mortality rate which is the variable of relevance.

Fig. 7. Quarterly occupied beds in English NHS hospitals and number of deaths (all-cause mortality) in England

Fig. 8. Estimated trend in deaths at Kings College Hospital (KCH) based on trends in the Local Authority areas from where patients are admitted

The trend for KCH was derived by adding together each Local Authority after weighting for the proportion of emergency admissions, Bromley 35%, Southwark 19%, Lambeth 18%, Lewisham 6%, Bexley 5%, Croydon 5%, Greenwich 3%

The second point of interest from Fig. A4 is the far higher volatility in the end-of-life component of costs in Milton Keynes compared to London. This reiterates the observation that some locations are subject to higher intrinsic volatility than others. The meaning of the saw-tooth patterns in deaths seen in Milton Keynes, and elsewhere throughout the UK, has been explained at great length elsewhere [14,31-35,40-41,49-50,62-65]. Yet, no government agency in the UK will acknowledge that these patterns even exist. Denial is clearly no basis for the rational planning of end-of-life services, beds and bed occupancy.

In conclusion, both the trend and the long-term volatility in bed demand arising from end-of-life will vary considerably between Local Authority locations and between hospitals servicing a variety of locations. Population demography will be totally unable to reflect these complex interactions.

7.2 Expected Growth in Deaths

Based on the 2012 subnational population projections prepared by the Office for National Statistics it was expected that deaths in England would increase by 17% in the 25-year interval 2013 to 2037. However, the median value at Local Authority level is +19%, and the interquartile range is 11% to 28%. Maximum expected increase is +55% in East Hampshire and +53% in Milton Keynes, and minimum expected is -9% in Norwich. A weighted composite of local authorities with highest numbers of emergency admissions at Kings College Hospital only gives +10% anticipated growth, i.e. less than the lower quartile.

As time progresses, different locations will experience significant shifts in their exposure to end-of-life bed demand and costs.

8. PRAGMATIC BED FORECASTING

The trends in admissions, especially those of a medical nature, are subject to a variety of forces, and as such the demographic component only represents a minimum case baseline increase [11,65]. The excellent review of Kendrick and Conway [8] discusses a wide range of societal factors which may lead to rising rates of admissions in elderly people. These can include people living alone, access to family and friends...
as unpaid care givers, risk aversion and care in nursing homes, etc.

A basic approach to forecasting will therefore always involve access to a trend from the past six to ten years, upon which various future scenarios are added, such as the anticipated increase in end-of-life related admissions and bed requirements plus a host of other societal and technological trends. The key point is to generate several scenarios with clearly stated assumptions. No process is perfect, and it is recognized that government policy, affordability, risk aversion will all play a major part in any final decision – even if all concerned recognize that it may well be a sub-optimal outcome.

This process can sadly be hijacked. The use of the Private Finance Initiative (PFI) in the UK and elsewhere, has led to high capital costs, and consequent emphasis on demographic-only forecasts to which were added wildly optimistic assumptions around future reductions in emergency admissions and length of stay, all to ‘predict’ the need for far smaller and ‘affordable’ hospitals [73-77]. This abuse of demographic forecasting, and its consequences (Fig. 1) serve as a salutary lesson to all.

There are several specific aspects of bed planning which can now be discussed.

8.1 Same Day Stay Admissions

In the early 1990’s health care systems around the world witnessed an increase in both elective and emergency same day stay admissions [78-81]. In the elective sphere these are referred to as ‘day surgery’ or ‘day case’ admissions and have arisen from substitution of otherwise overnight stay admissions due to advances in surgical technology such as key-hole surgery and new generation anesthetics. Elective admissions are therefore best forecast after combining same day and overnight stay admissions, and then making separate provision for the different theatres, day surgery units and overnight beds (with associated critical care beds) used to treat the respective patients.

However, a similar rise in same day emergency admissions has been driven by two factors. Firstly, the emergence of emergency observation/assessment wards where mainly elderly medical patients can receive rapid assessment, before either discharge back to the care of their GP, or further inpatient treatment. Rapid growth in the number of these units commences in the late 1990’s and there is ambiguity as to whether those ‘discharged’ on the same day are equivalent to an emergency department attendance [82-83]. In Australia, there was a 46% increase in same day emergency admissions between 1994 and 1999, accounting for 95% of the overall increase in medical emergency admissions [81].

In England (but not the rest of the UK), added impetus to the growth in same day emergency admissions came with the introduction in 2002 of a four-hour target for maximum waiting time in the emergency department [80-84]. This led to a rapid expansion in same day emergency admissions as hospitals opened assessment units as a means of circumventing the target [17,18]. During 2014/15 in England, there were some 1.6 million same day emergency admissions which accounted for 15% of orthopedic, 30% of medical and surgical, and 47% of pediatric emergency admissions [84].

There is huge variation (range 14% to 40%) in the proportion of same day emergency admissions between NHS hospitals in England [85], implying great ambiguity in the boundary between the emergency department and assessment units [86]. Due to a variety of factors the 4-hour target in England has come under increasing pressure, and between 2013/14 and 2015/16 same day stay emergency admissions to acute hospitals rose by 14.3% while emergency admissions with one or more days stay only increased by 3.3%. Clearly 14.3% growth in just two years has nothing to do with demographic change.

Any forecasting of emergency admissions (demographic-based or otherwise), should always treat same day admissions as a separate stream, and consequent calculation of same day bed numbers will depend on patterns of arrival throughout the day and average length of stay calculated in hours rather than days. In general, same day bed occupancy peaks around 2 pm in medical assessment/observation units.

8.2 Are Bed Numbers the Real Problem?

It has been recently demonstrated that the real problem may not be bed numbers per se, but the inability to rapidly flex staff numbers in response to volatile admissions [87]. The aim of health care cost containment is to reduce occupied bed-days, but to have a sufficient occupancy margin
to maintain efficient patient flow [87]. Beds are an asset for maintaining a highly efficient organization [88,89], however in the UK, an obsession with reducing bed numbers has led to dangerously high average occupancy (as per Fig. 1), with consequent queues to admission, and the hidden effects of high occupancy [87].

8.3 Incorporating Planned Reductions in Admissions

Some health care systems operate at a higher ratio of beds per death than others [90]. In the US, bed-days in the last six months of life ranged from 24 in New York to 11 in San Francisco [91]. One study investigating the anticipated 17% rise in deaths between 2012 and 2030 in the UK estimated that deaths at home were likely to fall from 31% to 18% with consequent pressure on hospital beds [92]. In England, all health and social care organizations have been mandated to work together to produce Sustainability and Transformation Plans (STPs) in a bid to reduce the reliance of the population on acute care [93].

The temptation is to produce a set of forecasts for large reductions in hospital bed numbers based on demographic forecasts plus planned reductions in admissions.

As stated above, the aim is to reduce occupied bed-days, but this does not automatically imply a reduction in bed numbers since additional beds are required to reduce the very high occupancy in Fig. 1 back to a safe and efficient level. It is in no one’s best interest to massage the numbers to produce a plan that is seemingly affordable (but only on paper).

9. ADMISSIONS PER DEATH

The ratio of admissions or bed-days per death has been proposed as a useful tool for tracking progress in reducing acute admissions via community-based alternatives [26,94-95]. Fig. 9 shows the number of consultant episodes (elective + emergency) per death in England in 2012/13 and 2013/14 for each of the age bands available in the Hospital Episode Statistics summary data [84]. As is expected the ratio of episodes per death (all-cause mortality) is very high in the early years of life since there are very few deaths.

The information in Fig. 9, does however, explain why the ratio of bed-days per death in London is so markedly higher than the rest of England [95]. London has a far higher proportion of elective bed days related to injury and ‘wear and tear’ occurring in a generally younger population with a far lower proportion of end-of-life care.

However, on this occasion an age-standardized measure of bed-days or admissions per death is a useful way of following if real shifts out of acute care are indeed occurring. The focus of attention may well be on tracking bed days per death in those aged 65+, and a suitable age-weighted approach would be justified.

Lastly, the concept that different communities experience markedly different proportions of beds devoted to end-of-life is illustrated in Fig. A 5 in the Appendix where the ratio of deaths per 1,000 population ranges from a minimum of 4 in London through to a maximum of 16 in rural and popular retirement destinations. Earlier comments regarding net migration out of London at retirement age are relevant in this context. The ratio of deaths per general practitioner (GP) shows even wider variation ranging from 7 to 27 (Fig. A6), indicating that disproportionate primary care resources are also required to deal with end-of-life care, especially in rural and retirement locations. This disproportionate level of resource input is also seen in the crude mortality rate of English hospitals (Fig. A7), which ranges from 1.3% to 5.4%.

Fig. A7 is important since it allows the estimations of the likely range in bed occupancy associated with end of life. From Fig. A7 there are 14 to 54 deaths per 1,000 acute-only admissions. In England, there are around 110 bed days per death (including mental health) [95] giving a maximum range of 4.2 to 16.3 beds per 1,000 admissions devoted to end-of-life care. Some of this is unpreventable, but an estimate of 15% which is preventable gives a ballpark figure of 0.6 to 2.4 beds per 1,000 (overnight stay) admissions as the potential reduction in the bed pool arising from a shift to a more palliative style of community-based care. See Section 10 for further discussion.

It is therefore clear that end-of-life plays a far greater role in bed demand in some locations than others. Modelling therefore needs to focus on beds required in the last year of life, versus beds required in those not in the last year of life. Certain diagnoses (such as presented in Table A2) will gain greater prominence in the end-of-life phase.
Consultant episodes in the 2012/13 and 2013/14 financial years in from Hospital Episode statistics [80], while deaths in the calendar years 2012 and 2013 are from the Office for National Statistics. Consultant episodes include both elective and emergency admissions, including same day stay admissions. Acute and mental health are included but maternity/midwifery are excluded. There can be more than one consultant episode per admission if the patient requires more complex care, although the ratio of consultant episodes per admission tends to be closer to 1.0 for elective admissions, except in Orthopedics if the patient requires convalescence in Geriatrics or Rehabilitation. An episode in an Assessment Unit followed by a move to care under a different consultant in General Medicine would count as two consultant episodes, etc. In rare cases (younger) patients can have over 15 consultant episodes per admission.

9.1 Roles for Migration and Deprivation

It should now be clear that it is where a person lives in the last year of life that drives acute (and other) resource consumption. Migration out of London after retirement has been highlighted as a contributing factor to the very young average age. Milton Keynes was also cited as a 1960’s new town where people are now relocating their elderly parents to provide a measure of end-of-life care. Indeed, popular retirement locations such as Devon have seen very little change in the total number of deaths over the past 15 years simply because the death of a retiree creates space for the arrival of another retiree. Frail elderly people are often moved to a limited number of nursing homes, where death makes space for the arrival of another frail resident.

In 1999 Brimblecombe et al. [96] noted that geographic variations in the age- and sex-standardized mortality rates could largely arise as an artefact of migration patterns. As an example, the presence of nursing homes leads to significant distortion in the age-standardized mortality rate and calculated life expectancy for small areas [97-98].

Hence areas such as London deliver care mainly devoted to children and diseases of middle age arising from ‘life in the fast lane’. The ethnic mix in London is likewise unique, and hence selective migration out of London at retirement is reserved for the wealthy, leaving a more deprived residue of the elderly subject to higher multi-morbidity [99], and hence patterns of bed utilization.

Deprivation is well known to be associated with higher age-standardized mortality due to death earlier in life. The utility of using the trend in deaths as a measure of end-of-life health care utilization is that no adjustment for deprivation is required. However, there do not appear to be any studies documenting whether end-of-life bed requirements are higher for persons from more deprived areas. The suspicion is that this may be
the case, and research is required to clarify this issue.

10. AVOIDING ‘FUTILE’ ACUTE INTERVENTION

This policy piece has identified that up to 55% of hospital (acute and mental health) bed days occur in the last year of life.

The period 22 weeks prior to death appears to mark a watershed for acute care [24]. It is at this point that acute intervention appears to become largely futile, including surgical and critical care interventions [100]. The emphasis at this point needs to shift to a palliative model for end-of-life care.

Supporting this proposal is a study from the USA investigating symptom trends in the last year of life over the interval 1998 to 2010. Over this interval, the proportion of persons experiencing ‘palliative’ symptoms such as moderate or severe pain, depression, confusion, dyspnea, incontinence, fatigue and anorexia all increased [101]. Interestingly a large increase in frequent vomiting and confusion by year of death was associated with the sudden death group, a trend upward in the prevalence of confusion was noted in the cancer death group, while the chronic lung disease and congestive heart failure group saw increases in depression and confusion over time. The frailty group showed increasing depression [101].

Likewise, a pan-European (SHARE) study demonstrated that almost all decedents received care in the last year of life, which was a crucial change compared to the previous year. In the previous year some 63% of persons had no limitations in the activities of daily living whereas in the last year some 58% had 1 or more limitations. Individuals dying of cancer were least impaired in the previous year and most so in the last year [102].

Is there any evidence to demonstrate that such large-scale reductions in acute bed days are possible? A study of hospital bed utilization between 2006/07 and 2009/10 in English Primary Care Organizations (PCOs) demonstrated that the Torbay integrated health and social care PCO operated at 33% fewer bed days per death after adjusting for the effects of deprivation, urban/rural mix and proportion of deaths outside hospital [90]. From this it should be clear that the figure of 55% of bed days at end-of-life is not an over-estimate. However, a word of caution is required since Torbay has 20 deaths per GP which is at the high end of Fig. A6, i.e. most other PCOs in England may be unable to achieve the potential reduction in bed days achieved in Torbay.

There are already plans to increase the degree of integration between health and social care in England via Sustainability and Transformation Plans (STPs) [93]. In an era of increasing medical litigation there is an even greater need for the development of tools based on blood biochemistry, vital signs, frailty and cognitive state [24,103-117], to enable acute and primary care physicians to make an informed (and documented) decision as to when to initiate this transition.

Hospital doctors, especially geriatricians, will become even more central in this transition. Suitably qualified community nurses and health care assistants will need to be trained, and seamless information sharing will become the norm. Islands of biochemical data currently held in hospital and GP data bases will need to be accessible over the life time of a patient to track health progression and help determine the point at which health state has moved to end-of-life care [24].

11. CONCLUSIONS

Demographic forecasting is a simplistic tool subject to considerable underestimation of future activity (and beds), especially for medical conditions which fall into multi-morbidity clusters. End-of-life may account for up to 55% of a person’s lifetime hospital bed utilization, and the expected trends in deaths need to be incorporated into forecasting methodologies. There is wide variation across England in the trends in death and the relative importance of end-of-life upon costs and bed utilization, and related primary and social care resources. Of all locations, parts of London have experienced the greatest decline in deaths over the past 15 years, and are least influenced by end-of-life care requirements. The Kings College hospital in London is an example of a hospital with a nationally low contribution from end-of-life, plus very low anticipated future growth in deaths in the population surrounding the hospital sites. While London accounts for 15% of the population of England, it only accounts for 10% of the deaths. London-wide schemes for end of life care are probably required. Rural areas with the
highest burden of end-of-life care will require additional funding to compensate for the greater distances involved in providing home-based care.

Reliance on the Public Finance Initiative in the UK emphasized the use of demographic forecasting simply because it gave answers which were conveniently lower than reality, but could be used to ‘justify’ building smaller hospitals. How hospitals are run appears to require a paradigm shift [88], to account for the flexible staffing implied by high volatility in bed demand induced by seasonal, and infectious pressures [87].

Unexpected trends in hospital bed demand are more the norm than the exception [89]. An inconvenient fact which seems to have been largely ignored. Indeed, the year-to-year volatility in deaths at local authority area is far higher than any government department appears willing to acknowledge [67-72], and implies far higher flexibility in the allocation of funds at local level.

The very same volatility seen in deaths and acute bed demand will place equal pressures on all community-based schemes for admission avoidance, and if overwhelmed, then acute beds will once again become the default position.

A generation of models need to be urgently developed which augment basic demographic trends with additional relevant societal, technological and disease incidence trends, plus a stream devoted to the last year of life.

CONSENT
It is not applicable.

ETHICAL APPROVAL
It is not applicable.

ACKNOWLEDGEMENTS
Discussion with Dr John Kellett and the suggestions of the BJMMR reviewers are acknowledged with gratitude.

COMPETING INTERESTS
Authors have declared that no competing interests exist.

REFERENCES


63. Jones R. Simulated rectangular wave infectious-like events replicate the diversity of time-profiles observed in real-world running 12 month totals of admissions or deaths. FGNAMB. 2015;1(3):78-79. DOI: 10.15761/FGNAMB.1000114


65. Jones R. Rising emergency admissions in the UK and the elephant in the room.


## APPENDIX

Table A1. Top 25 diagnoses relating to infants in their first year of life, 2012/13 England

<table>
<thead>
<tr>
<th>Diagnosis/Condition</th>
<th>Proportion of admissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short gestation and low birth weight</td>
<td>9.6%</td>
</tr>
<tr>
<td>Neonatal jaundice</td>
<td>6.2%</td>
</tr>
<tr>
<td>Acute bronchiolitis</td>
<td>6.0%</td>
</tr>
<tr>
<td>Intrauterine hypoxia</td>
<td>4.7%</td>
</tr>
<tr>
<td>Medical observation and evaluation</td>
<td>3.7%</td>
</tr>
<tr>
<td>Respiratory distress of new born</td>
<td>3.4%</td>
</tr>
<tr>
<td>Viral infection of unspecified site</td>
<td>3.3%</td>
</tr>
<tr>
<td>Long gestation and high birth weight</td>
<td>3.0%</td>
</tr>
<tr>
<td>Other perinatal respiratory conditions</td>
<td>2.9%</td>
</tr>
<tr>
<td>Feeding problems of new born</td>
<td>2.9%</td>
</tr>
<tr>
<td>Acute upper respiratory infections of multiple and unspecified sites</td>
<td>2.8%</td>
</tr>
<tr>
<td>Slow fetal growth and fetal malnutrition</td>
<td>2.0%</td>
</tr>
<tr>
<td>Viral and other specified intestinal infections</td>
<td>1.7%</td>
</tr>
<tr>
<td>Persons encountering health services in other circumstances</td>
<td>1.7%</td>
</tr>
<tr>
<td>Disorders of carbohydrate metabolism specific to fetus and new born</td>
<td>1.4%</td>
</tr>
<tr>
<td>Other gastroenteritis and colitis of infectious and unspecified origin</td>
<td>1.3%</td>
</tr>
<tr>
<td>Gastro-oesophageal reflux disease</td>
<td>1.2%</td>
</tr>
<tr>
<td>Bacterial sepsis of new born</td>
<td>1.2%</td>
</tr>
<tr>
<td>Other conditions of integument specific to fetus and new born</td>
<td>1.1%</td>
</tr>
<tr>
<td>Other conditions originating in the perinatal period</td>
<td>1.1%</td>
</tr>
<tr>
<td>Abnormalities of breathing</td>
<td>1.1%</td>
</tr>
<tr>
<td>Other disorders of urinary system</td>
<td>1.0%</td>
</tr>
<tr>
<td>Other congenital malformations of tongue, mouth and pharynx</td>
<td>0.9%</td>
</tr>
<tr>
<td>Fever of unknown origin</td>
<td>0.9%</td>
</tr>
<tr>
<td>Unspecified acute lower respiratory infection</td>
<td>0.8%</td>
</tr>
</tbody>
</table>

Table A2. Top 25 diagnoses relating to patients aged 90+, 2012/13 England

<table>
<thead>
<tr>
<th>Diagnosis/Condition</th>
<th>Proportion of admissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pneumonia, organism unspecified</td>
<td>18.2%</td>
</tr>
<tr>
<td>Other disorders of urinary system</td>
<td>15.5%</td>
</tr>
<tr>
<td>Fracture of femur</td>
<td>8.5%</td>
</tr>
<tr>
<td>Symptoms involving the nervous and musculoskeletal systems</td>
<td>6.9%</td>
</tr>
<tr>
<td>Heart failure</td>
<td>6.9%</td>
</tr>
<tr>
<td>Unspecified acute lower respiratory infection</td>
<td>6.4%</td>
</tr>
<tr>
<td>Cerebral infarction</td>
<td>5.1%</td>
</tr>
<tr>
<td>Acute myocardial infarction</td>
<td>3.8%</td>
</tr>
<tr>
<td>Syncope and collapse</td>
<td>3.7%</td>
</tr>
<tr>
<td>Other chronic obstructive pulmonary disease</td>
<td>3.1%</td>
</tr>
<tr>
<td>Other and senile cataract</td>
<td>5.2%</td>
</tr>
<tr>
<td>Acute renal failure</td>
<td>3.0%</td>
</tr>
<tr>
<td>Other retinal disorders</td>
<td>2.9%</td>
</tr>
<tr>
<td>Atrial fibrillation and flutter</td>
<td>2.9%</td>
</tr>
<tr>
<td>Other diseases of digestive system</td>
<td>2.9%</td>
</tr>
<tr>
<td>Pain in throat and chest</td>
<td>2.8%</td>
</tr>
<tr>
<td>Other malignant neoplasms of skin</td>
<td>2.7%</td>
</tr>
<tr>
<td>Symptoms and signs involving cognitive functions and awareness</td>
<td>2.7%</td>
</tr>
<tr>
<td>Open wound of head</td>
<td>2.7%</td>
</tr>
<tr>
<td>Cellulitis</td>
<td>2.3%</td>
</tr>
<tr>
<td>Other gastroenteritis and colitis of infectious and unspecified origin</td>
<td>2.3%</td>
</tr>
<tr>
<td>Unknown and unspecified causes of morbidity</td>
<td>2.2%</td>
</tr>
<tr>
<td>Other joint disorders, not elsewhere classified</td>
<td>2.0%</td>
</tr>
<tr>
<td>Superficial injury of head</td>
<td>2.0%</td>
</tr>
<tr>
<td>Pneumonitis due to solids and liquids</td>
<td>1.9%</td>
</tr>
</tbody>
</table>
Fig. A1. Deaths per 1,000 population by single-year-of-age in England & Wales (2015), by gender
Both deaths and population in England & Wales 2015 are from the Office for National statistics. The disparity in deaths between 16 to 24 is largely due to adolescent males engaging in high risk behaviors. The last data point covers ages 90 and above. Extrapolation to 1,000 deaths per 1,000 population marks the current maximum age at death in England and Wales.

Fig. A2. Quarterly average occupied beds (midnight) per 100 deaths (all-cause mortality) in England, average 2010 to 2016
As is implied from Fig. 5, there has been an ongoing reduction in the ratio of occupied beds per 100 deaths over time. Hence between 2010 and 2016 the ratio has decreased by the following amounts in each quarter, 0.9 per year (December), 1.4 per year (September), 1.5 per year March/June. There is some evidence that this reduction may be declining over time, i.e. the trend is not linear.
Fig. A3. Adjusted difference for admissions in 2009/10 (Swine flu) versus 2008/09 (without Swine flu)

To adjust for the underlying growth in admissions, the difference between 2009/10 and 2008/09 was compared to the difference between 2008/09 and 2007/08 (assumed underlying growth). An additional adjustment has been applied to estimate the impact of the growth in same day stay admissions during a time when medical assessment units were being opened in various hospitals throughout England.

Fig. A4. Trend in deaths for Inner London compared with Milton Keynes

Monthly deaths by district of residence is from the Office for National Statistics.
Fig. A5. Ratio of deaths per 1,000 population for Local Authority populations in England and Wales, 2015

Calendar year deaths and population by district of residence is from the Office for National Statistics.

Fig. A6. Ratio of deaths per whole time equivalent (WTE) General Practitioners (GPs) in English Clinical Commissioning Groups.

Deaths in English Clinical Commissioning Groups are from http://www.endoflife-care-intelligence.org.uk/profiles/CCGs/Place_of_Death/atlas.html while the number of General Practitioners (WTE) is from NHS Digital http://content.digital.nhs.uk/catalogue/PUB21772
Fig. A7. Crude mortality rate (including deaths within 30 days of discharge) for English hospitals

Data is from the Summary Hospital Mortality Index (SHMI) found on the NHS Digital website http://content.digital.nhs.uk/SHMI. The crude mortality rate is an average between 2013 and 2015.

© 2017 Beeknoo and Jones; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/17161