‘Chaplains for well-being’ in primary care: analysis of the results of a retrospective study

Peter Kevern1 and Lisa Hill2

1Senior Lecturer, Faculty of Health Sciences, Staffordshire University, Stafford, UK
2Senior Commissioning Manager, Sandwell and West Birmingham CCG, Stafford, UK

Aim: To analyse quantitative changes in patient well-being concurrent with chaplaincy interventions in a retrospective study of a group of Primary Care centres in Sandwell and West Birmingham, United Kingdom. Background: Anecdotal evidence suggests that support from trained Primary Care Chaplains may be particularly useful for those with subclinical mental health issues; it can reduce the tendency to ‘medicalise unhappiness’ and is a positive response to patients with medically unexplained symptoms. However, to date there has been no published research attempting to quantify their contribution.

Method: Data were gathered from a group of Primary Care Centres, which make use of a shared Chaplaincy service. Demographic data and pre–post scores on the Warwick and Edinburgh Mental Wellbeing Scale (WEMWBS) were collected for patients who had attended consultations with a Chaplain. These were subjected to tests of statistical significance to evaluate the possible contribution of chaplaincy to patient well-being along with possible confounding variables. Findings: a substantial improvement in WEMWBS scores (mean = 9 points, BCa 95% CI [7.23, 10.79], P = 0.001) post-intervention. The improvement in scores was highest for those with initially lower levels of well-being. There is therefore evidence that chaplaincy interventions correlate with an improvement of holistic well-being as measured by a WEMWBS score. A prospective study on a larger scale would provide more detailed information on the interaction of possible variables. Further study is also required to evaluate the implications of this result for patient outcomes and GP resources. The efficacy of Primary Care Chaplaincy is under-researched and difficult to measure. This paper represents the first attempt to quantify a measurable improvement in the well-being of patients who are referred to the service.

Key words: Chaplain; general practice; mental health; primary care; wellbeing; WEMWBS

Received 28 May 2013; revised 25 September 2013; first published online 22 January 2014

Introduction

In principle, ‘spiritual care’ is now widely accepted to have a place in health care in the United Kingdom (Department of Health (DoH), 2003; Mental Health Foundation, 2007; Care Services Improvement Partnership, 2008; DoH, 2009; NHS Education for Scotland (EfS), 2009a, 2009b). However, there is a wide variation in the way this principle has been applied to different fields of health care. At one extreme, spiritual care has been enthusiastically adopted and developed in relation to palliative care and mental health (Royal College of Psychiatrists, 2006; National Council for Palliative Care, 2011); whereas much less progress has been made in acute hospital care (Ronaldson et al., 2012).

On reflection, this is hardly surprising. Although definitions vary, in general spiritual care assumes a concern for the whole person, for their holistic well-being and relationships over time, as illustrated in this description from NHS Education for Scotland:

Spiritual Care is that care which recognises and responds to the needs of the human spirit...
when faced with trauma, ill health or sadness and can include the need for meaning, for self-worth, to express oneself, for faith support, perhaps for rites or prayer or sacrament, or simply for a sensitive listener (EfS, 2009a: 6).

As such, it is unlikely to be of much value in disciplines such as surgery (Taylor et al., 2011). However, one might expect notions of spiritual care to have some appeal to general practitioners (GPs) given that they are ‘generalists’ and therefore in some way committed to an holistic view of health care. It is certainly the case that some of them recognise a ‘spiritual’ dimension to patients’ well-being, ‘a need to love and be loved; a need to feel worthwhile; and a need for meaning and purpose’ (Bryson et al., 2012: 16). Many more recognise the importance of continuing relationships, particularly in relation to depression and low-level chronic illness (Cocksedge and May, 2005). There is also significant evidence (albeit mostly from the different religious ecology of the United States of America) that patients would like to discuss their spirituality, religion, values or matters of existential concern in the context of primary care (Brush and McGee, 1999; Wilson et al., 2000; MacLean et al., 2003; Ellis and Campbell, 2004; McCord et al., 2004).

However, there are barriers to GPs offering spiritual care. Good spiritual care is demanding: it requires time, skills and resources GPs may not have (Kliewer, 2004; Tanyi et al., 2009) and is difficult to contain within boundaries (Cocksedge and May, 2009). Many GPs are uncomfortable discussing spiritual needs and uncertain of patients’ responses (Brush and Daly, 2000). Consequently, it does not receive the attention many GPs feel it should have (Vermandere et al., 2011).

It is reasonable to infer that GPs themselves may not be the best people to offer spiritual care, and one alternative is for them to use ‘Chaplains’ who may offer it on their behalf. However, there is very little evidence regarding the potential and pitfalls of Primary Care Chaplaincy services. Little is known about what such services might achieve, how Chaplains might be trained and supervised, or how they might be viewed by service users. This is partly a consequence of the way such chaplaincy services have been set up (typically by a single practice, on a small scale and for a limited period) but also of the nature of chaplaincy (which, by definition, resists reduction to measurable outcomes).

Only two schemes have been subjected to analysis to date, and in both cases the analysis is at an early stage. The Community Chaplaincy Listening Service is being developed by NHS Education for Scotland. It represents a collaboration between eight health boards and makes use of 15 Hospital Chaplains who have agreed to attend 18 GP practices on an occasional or part-time basis. A distinctive feature of this initiative is that, from the start, it was intended to produce a model that could be applied across Scotland. It therefore began life in 2010 as a theoretical construct of chaplaincy as centred on the therapeutic and existential importance of ‘listening’ (Mowat et al., 2013). The resultant service is now in the process of phase 2 evaluation. So far, primarily qualitative data have been gathered and analysed (Mowat et al., 2012; Bunniss et al., 2013), but a PROM (Patient-Recorded Outcome Measure) is in the process of development and testing (Bunniss et al., 2013; Snowden et al., 2013a; 2013b). In contrast to the top-down design of the Scottish model, the Sandwell and West Birmingham CCG Chaplains for Wellbeing Service derived from the work of one GP practice, Karis Medical Centre. There are currently five Chaplains, employed as part of the Sandwell ‘hub’, meaning that they can receive referrals from all 134 Primary care centres in the region (Bryson et al., 2012). In this model there has never been a precisely articulated version of the role of a Chaplain, although a list of eight key functions has emerged by a form of consensus:

... the combination of eight key functions that makes the Community Healthcare Chaplain truly distinctive: listening, compassionate presence, facilitating the search for meaning, discerning the signs of life, offering appropriate ritual, offering prayer, providing support in death and dying, and pastoral care of staff (Bryson et al., 2012, p. 20).

Sandwell is an urban area of the West Midlands, the fourth most deprived local authority in England in 2011 and the bottom quartile of local authorities across England for social deprivation. It has a practice-based population of 310,000 and high levels both of ethnic diversity and of deprivation. A total of 23% of the population are from a recognised...
minority ethnic group (mostly of Pakistani and Indian origin). The area has high levels of ‘non-decent’ homes (Middleton, 2010). The Sandwell ‘hub’ was set up to address some of the chronic and low-level threats to well-being across the region with a range of available interventions such as user-led groups, psychotherapy and welfare rights services as well as the Chaplains for Wellbeing Service (Das, 2012).

This service has the potential to make a distinctive contribution to the available knowledge-base on Primary Care Chaplaincy for three interrelated reasons: its relative longevity; the distinctive features (high-population density and diversity) of the area it serves; and the way it has evolved in response to local needs. Initially, the constant development and change in the service made it difficult to establish a sufficiently firm evidence base. However, with the introduction of standardised record-keeping through the ‘Sandwell Hub’ and the adoption of WEMWBS (Warwick and Edinburgh Mental Wellbeing Scale) it has now become possible to quantify the impact of chaplaincy services.

**Method**

A period was identified (1 January 2011–1 January 2013) in which chaplaincy provision was stable, being offered by the same two Chaplains and with patients’ well-being assessed against the WEMWBS scale. The advantages of the WEMWBS scale were that it measures ‘mental well-being’ rather than a contested and theologically loaded concept such as ‘spirituality’, and that it was designed to measure overall well-being rather than a measure mental illness (which tends to overlook changes in the mental well-being of the ‘healthy’ population, Stewart-Brown and Janmohamed, 2008, p.1). The scale shows broad reliability and validity (Tennant et al., 2007; Stewart-Brown et al., 2009; Maheswaran et al., 2012) as well as having been tested cross-culturally (Taggart et al., 2013), a significant consideration in Sandwell.

Throughout, WEMWBS was administered by the Chaplains themselves who would normally offer an initial assessment, and one before closing.
the patient record. However, it transpired that many patients who attended at least one session with the Chaplain did not have two WEMWBS scores, for a variety of reasons. In many cases, the patient attended only once, particularly those attending the Primary Care Walk-in Centre that supplied a proportion of the data. In others, contact was lost (often after several missed appointments) before a second measurement of wellbeing could be taken.

Analysis was conducted in IBM SPSS-21. The core dataset for purposes of analysis was those patients who had ‘Pre and Post’ scores, that is, with at least two WEMWBS assessments (n = 107). For these patients, data were compiled from the patient record for age; sex; employment; ethnicity; Chaplain; WEMWBS score (first, last); total number of visits; and number of visits between first and last scores. However, this represented only a sample of the population of attendees (N = 246) for which data were less complete.

The analysis of the data therefore required the testing of three hypotheses:

1. The sample (n) is representative of the total population (N) accessing the chaplaincy service during this period.
2. There will be a significant change in WEMWBS scores post-intervention.
3. Changes in WEMWBS score are evenly distributed across sub-groups of the sample as recorded in the demographic data.

Results

Comparison between population and sample data

χ²-tests were conducted to explore association between population/sample and variables as follows:

(a) Sex: proportions of M:F were compared for Population (Predictor) and Sample (sample). χ²(1) = 0.101, P = 0.75.
(b) **Ethnicity:** data were necessarily approximate, depending as they did on self-report. They were grouped as White/White British, Black/Black British, Asian/Asian British, Other/No Record $\chi^2(3) = 0.039$, $P = 0.998$.

(c) **Employment:** data were grouped as Unemployed, Employed, No Record $\chi^2(3) = 2.151$, $P = 0.34$. Although not significant in itself, this reflected a slight increase in the proportion in employment in the Sample ($n$) compared with the Population ($N$).

Parametric tests were conducted on the following variables. As Kolmogorov–Smirnov tests showed significant departures from a normal distribution for the data on patient Age and Initial WEMWBS score, data were bootstrapped (1000 sample, 95% CI) for parametric tests.

(d) **Age:** mean ages for Population ($N$) and Sample ($n$) were compared using an independent-samples two-tailed $t$-test (bootstrapped). For Population, $M = 44.49$, SD = 15.04, SE = 0.99. For Sample, $M = 42.79$, SD = 14.60, SE = 1.53 ($F = 0.001$, significance = 0.98). Mean difference = 1.7, BCa 95% CI $[−1.9, 5.19]$. This difference was not significant, $t(321) = 0.93$, $P = 0.38$.

(e) WEMWBS initial scores for Population ($N$) and Sample ($n$) were compared using an independent-samples two-tailed $t$-test (bootstrapped). For Population, $M = 34.95$, SD = 10.78, SE = 0.73. For Sample, $M = 35.72$, SD = 11.22, SE = 1.08 ($F = 0.12$, significance = 0.73). Mean difference = 0.78, BCa 95% CI $[−3.47 – 1.68]$. This difference was not significant, $t(324) = 0.60$, $P = 0.55$.

**Summary**

Analysis indicates that there is a close correlation between the characteristics of the sample.

---

**Figure 3** Comparison of population and sample distribution by ethnicity

[Diagram showing distribution by ethnicity]
and those of the population from which it is drawn. Therefore, data from the sample can be extrapolated with a high degree of confidence for the total population of individuals accessing the Sandwell Chaplains for Wellbeing Service in the period under examination. There are a few minor variations – notably in the age and employment status of those who go on to receive a second or subsequent WEMWBS score, and these may repay examination in a larger study.

**Test for presence or absence of significant change in WEMWBS scores for service users**

The pre- and post-scores on the WEMWBS scale were compared for the sample \( (n = 107) \) with the following results (theoretical maximum range = 14–70):

- **Pre-score**: \( M = 35.72, \) SD 11.217, BCa 95% CI [33.60, 38.00], median = 34.
- **Post-score**: \( M = 44.76, \) SD 10.624, BCa 95% CI [42.74, 46.78], median = 46.

Paired samples \( t \)-test (two tailed, bootstrapped), mean difference = 9.04, SE = 0.91, BCa 95% CI [7.23, 10.79] was significant, \( t(106) = 9.62, P < 0.001 \). Cohen’s \( d = 0.85 \), representing a large effect.

As the median scores represented a larger difference (12 points) than the mean scores (9 points), the former were tested for significance with the Related-Samples Wilcoxon Signed-Rank Test (two sided). \( T = 5192, P < 0.001, r = 0.72 \), representing a large effect.

**Summary**

A mean improvement of 9 points on the WEMWBS scale was observed for users of the ‘Chaplains for Wellbeing’ service who have two successive measurements on the scale. This is significant at the \( P \leq 0.001 \) level.
In their important discussion of the topic, Maheswaran et al. (2012) conclude that a clinically 'meaningful' change in WEMWBS score is between 1.0 and 2.77 SEM, which is generally between 3 and 8 points. However, in terms of an individual's sense of well-being a more helpful measure might be derived from the suggestion of Jaeschke et al. (1989) that an improvement of 0.5 on each item in a 5-point Likert scale represents an improvement that an individual would deem important. This suggests an improvement of 7 points across the 14 items would be of importance to the individual concerned. On either estimation, the reported improvement of about 9 points is worthy of note.

In addition, the fact that the median increased by more than the mean when the two samples were compared gives an indication that the improvement in scores is asymmetrically distributed, with a greater improvement encountered in those presenting with a lower initial score. This hypothesis will be tested in the following section.

**Tests of the evenness of distribution of WEMWBS scores across sub-groups**

One-way ANOVA was conducted to compare means for first and last WEMWBS scores, and the difference between them (Last – First) for Sex, Employment status and Ethnicity. Results are summarised in Table 1 and in Figures 6–8.

(a) **Sex**: no significant variation between the scores for women and men.

(b) **Employment**: the variation in the final WEMWBS scores approaches the significance threshold. This reflects the fact that the mean score for those in employment (47.87) is higher than the total mean (44.89).

(c) **Ethnicity**: the variation in initial WEMWBS score is significant at the $P=0.01$ level. This reflects a mean score for the Black British category (44.86) substantially higher than the total mean (35.72).

*Primary Health Care Research and Development* 2015; **16**: 87–99
Parametric tests were conducted on the following:

(d) Age: Pearson’s test (two tailed) was used to investigate a possible correlation between age and change in WEMWBS score. \( N = 103, \ r = 0.13, \ P = 0.21. \) Bootstrapped 95\% CI \([-0.11, 0.35].\)

(e) Number of visits: Pearson’s test (two tailed) was used to investigate a possible correlation between number of visits between successive

### Table 1 Results of one-way ANOVA for initial (Pre), final (Post) and final – initial (Post – Pre) WEMWBS mean scores

<table>
<thead>
<tr>
<th>Factor: sex (female, male)</th>
<th>Sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>WEMWBS pre (session)</td>
<td>125.981</td>
<td>1</td>
<td>125.981</td>
<td>1.001</td>
<td>0.319</td>
</tr>
<tr>
<td>WEMWBS post (session)</td>
<td>0.016</td>
<td>1</td>
<td>0.016</td>
<td>0.000</td>
<td>0.991</td>
</tr>
<tr>
<td>Post – Pre</td>
<td>118.802</td>
<td>1</td>
<td>118.802</td>
<td>1.269</td>
<td>0.263</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor: Employment (Unemployed, Employed, No Record)</th>
<th>Sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>WEMWBS pre (session)</td>
<td>85.602</td>
<td>2</td>
<td>42.801</td>
<td>0.333</td>
<td>0.717</td>
</tr>
<tr>
<td>WEMWBS post (session)</td>
<td>527.377</td>
<td>2</td>
<td>263.688</td>
<td>2.415</td>
<td>0.094</td>
</tr>
<tr>
<td>Post – Pre</td>
<td>231.864</td>
<td>2</td>
<td>115.932</td>
<td>1.237</td>
<td>0.294</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor: Ethnicity (White British, Black British, Asian British, Other, No Record)</th>
<th>Sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>WEMWBS pre (session)</td>
<td>1646.661</td>
<td>4</td>
<td>411.665</td>
<td>3.592</td>
<td>0.009</td>
</tr>
<tr>
<td>WEMWBS post (session)</td>
<td>398.170</td>
<td>4</td>
<td>99.542</td>
<td>0.878</td>
<td>0.480</td>
</tr>
<tr>
<td>Post – Pre</td>
<td>608.204</td>
<td>4</td>
<td>152.051</td>
<td>1.660</td>
<td>0.165</td>
</tr>
</tbody>
</table>

WEMWBS = Warwick and Edinburgh Mental Wellbeing Scale.

**Figure 6** Comparison of mean WEMWBS scores (pre and post) by sex. WEMWBS = Warwick and Edinburgh Mental Wellbeing Scale.
WEMWBS scores and change in WEMWBS score. \( N = 107, \ r = 0.022, \ P = 0.82 \). Bootstrapped 95% CI \([-0.18, 0.20]\).

(f) Initial WEMWBS score and improvement in score: noting the discrepancy between mean and median improvement [see (2) above], Pearson’s test (two tailed) was used to investigate a possible correlation between initial WEMWBS score and change in WEMWBS score. \( N = 107, \ r = -0.481, \ P < 0.001 \). Bootstrapped 95% CI \([-0.614, -0.331]\), \( R^2 = 0.23 \) (Figure 9).

**Summary**

There was no evidence that either the initial or final WEMWBS score was dependent on the variables Sex, Age or the Number of Visits. The contribution of Employment status was less clear, though falling short of the threshold for significance. Ethnicity appeared to have a significant effect on the initial WEMWBS score (and therefore may reflect uneven recruitment to the service), but the effect was not reproduced in the final score.

A statistically significant inverse relationship was identified between initial WEMWBS score and the size of the improvement, which accounted for ~23% of the variance in the improvement (see Figure 2).

**Discussion**

Data on the outcomes of a chaplaincy intervention are inherently problematic, for a range of reasons. In the first place, there is no clear ‘trigger’ for referral or self-referral to the chaplaincy service, so a range of patients may be expected. The range of possible variables in response to the service (eg, age, sex, religion, ethnicity, number of sessions attended) introduces a high variability in

Primary Health Care Research and Development 2015; 16: 87–99
small samples, which would not necessarily be reduced by a larger sample drawn over a wider area or longer time period. Finally, the employment of the WEMWBS tool is not entirely consistent within the service: Chaplains exercise some discretion about when it is appropriate to introduce the tool because a patient may arrive at the initial meeting in acute distress. However, this variation in the use of the measure is as likely to add to the reliability of the data rather than detract from it, as it reduces the danger of a falsely low initial reading reflecting a transient fluctuation in well-being.

Given this potential variability, the consistency of the findings in this study is worthy of note. Analysis of the available data from two years of the operation of the Sandwell Chaplains for Wellbeing Service indicates that:

1. Although partial, the available data can reasonably be assumed to provide an accurate picture of the demography and outcomes of the population accessing the service.
2. There is a significant mean improvement on the WEMWBS measure of well-being for service users attending more than once. This does not in itself provide conclusive evidence for the effectiveness of the service, because the data can take no account of some potential confounding variables (such as the possibility that service users attend when they are already beginning to improve). However, it indicates the potential value of a prospective intervention study.
3. There is no evidence for dependency of this improvement on Sex, Ethnicity, Age or Number of Visits. There is some suggestion that Employment status may affect the size of the improvement, and this may repay further study. The main independent variable influencing the size of the improvement appears to be the initial WEMWBS score, with which

**Figure 8** Comparison of mean WEMWBS scores (pre and post) by ethnicity. WEMWBS = Warwick and Edinburgh Mental Wellbeing Scale.
there is a statistically significant negative correlation that accounts for 23% of the total variance in improvement. Some caution needs to be inserted, because the result may also in part be an artefact of the system of measurement – a ‘regression to the mean’ by those at the statistical extremes – but the result implies that those with the lowest levels of well-being stand to benefit the most from the service.

Conclusions

This study provides initial evidence that the provision of a ‘Chaplains for Wellbeing’ service in primary care improves mental health and well-being among those referred to it, as measured on the WEMWBS scale. This finding corroborates qualitative and anecdotal findings in the earlier pilot (Bryson et al., 2012) and the concurrent Scottish study (Mowat et al., 2013) that report an improvement in well-being as reported by service users.

The limitations on this conclusion are to a large extent the perennial ones that bedevil retrospective studies, and particularly retrospective population studies. It is not possible to exclude all potential confounding variables, to randomise the sample or to compare against a control, so the effect of (for example) self-selection by those already beginning to improve cannot be estimated. For these reasons, there is a need for a larger, prospective intervention study before the robustness of the findings can be assessed.

Further work also needs to be undertaken to establish whether this effect is maintained or enhanced as the service grows (it has recently engaged three more Chaplains); and the extent to which a similar effect is demonstrated whenever patients are given time to talk freely and be listened to, regardless of the label or function of the service they access.

Finally, if the chaplaincy service is to be deemed worthy of commissioning across a wider area, there is also a need to establish the effects of this improvement in WEMWBS score on key indicators of health such as frequency of

Figure 9  Correlation between initial WEMWBS score and number of points improvement. WEMWBS = Warwick and Edinburgh Mental Wellbeing Scale.
attendance at the GP’s surgery and the rate of antidepressant use.

Acknowledgements

The authors thank Dr Ross Bryson, Rev Fiona Collins and Mrs Rita Gupta for comments and support; they also thank Professor Tony Stewart and Professor David Clark-Carter for advice on the statistical analysis; and the reviewers of an earlier draft of this paper for their valuable comments and advice.

Financial Support

Funding for this study was provided by Sandwell and West Birmingham CCG.

Conflicts of Interest

The second author is employed by Sandwell and West Birmingham CCG.

Ethical Standards

This is a secondary analysis of data that was originally gathered as part of a service evaluation, and did not require any additional intervention.

Data were accessed in an anonymous form through the electronic records according to Caldicott guidelines.

References


Primary Health Care Research and Development 2015; 16: 87–99


Department of Health. 2003: NHS Chaplaincy: meeting the religious and spiritual needs of patients and staff. London: DoH.


NHS Education for Scotland. 2009b: Spiritual and religious care capabilities and competences for healthcare Chaplains.


