Journal of Evaluation in Clinical Practice ISSN 1356-1294

An independent evaluation of the modernization of NHS endoscopy services in England: data poverty and no improvement

Kymberley Thorne PhD MPhil BSc,[1] Hayley A. Hutchings PhD FHEA BSc[2] and Glyn Elwyn PhD FRCGP MSc BCh MBBA[3]

[1]Health Services Researcher, [2]Senior Lecturer, Centre for Health Information, Research and Evaluation (CHIRAL), School of Medicine, Swansea University, Swansea, UK
[3]Professor, Department of Primary Care and Public Health, School of Medicine, Cardiff University, Cardiff, UK

Keywords
activity, data availability, data collection, endoscopy, modernization, service evaluation

Correspondence
Dr Kymberley Thorne
Centre for Health Information
Research and Evaluation (CHIRAL)
School of Medicine
Swansea University
Singleton Park
Swansea SA2 8PP
UK
E-mail: k.thorne@swansea.ac.uk

Accepted for publication: 22 January 2009

doi:10.1111/j.1365-2753.2009.01177.x

Abstract

Rationale, aims and objectives The Modernising Endoscopy Services (MES) programme introduced a focussed modernization drive and data collection regime to English NHS endoscopy services. We independently evaluated the MES programme by comparing routinely collected, service-related endoscopy data from sites that participated in the MES programme and sites that did not.

Methods A random selection of 10 endoscopy units who had participated in the MES programme (intervention sites) were compared with a random selection of 10 endoscopy units who redesigned their services independently (control sites). Data on demand, numbers waiting, activity and cancellations were collected for eight time points between January 2003 and April 2006. Data were aggregated into intervention and control groups for statistical analysis using a two-way analysis of variance. Activity data were validated using an equivalent Hospital Episode Statistics dataset.

Results Data were not routinely collected by 11 of 19 endoscopy units. Trust-held datasets were subsequently included to address problems with data availability. The accuracy of the Activity data was successfully validated. Statistical analysis of the data showed that neither the intervention group nor the control group were able to significantly improve their services over time. There was also no significant difference between the intervention group and the control group in the improvement of their endoscopy services at any point time.

Conclusions Based on the data collected, the intervention programme did not significantly improve NHS endoscopy services in England over and above what could have been achieved independently with only the intention to redesign.

Introduction

Since the publication of the NHS Plan [1] and NHS Cancer Plan [2] in 2000, NHS endoscopy services have been under consistent pressure to modernize within a strict budget from a patient-centred perspective as the Government imposed increasingly shorter waiting time targets. There were also changes to Department of Health guidelines to make endoscopy the preferred diagnostic procedure and the introduction of the Two Week Rule to improve the outcome for patients with suspected cancers. All of these factors contributed to the need to improve NHS endoscopy services in a novel but sustainable and cost-effective manner.

The NHS Modernisation Agency (NHSMA) was established in 2001 to help facilitate the redesign of a range of NHS services in England, one of which was the Modernising Endoscopy Services (MES) programme. This programme aimed to modernize NHS endoscopy services in England with the patient at its centre, to implement booking and choice and to demonstrate that improvements could be made by a systematic approach to service redesign [3]. Endoscopy units in England were asked to submit an application form describing their current services and their plans for redesign to the NHSMA. From a total of 99 endoscopy units, only 26 were successful in their application and they were subsequently given £30 000 to fund the NHSMA-approved redesign plans from their bids over a 12-month period. The MES programme sites were also expected to complete a data collection tool called the Toolkit™, which required the daily input of referral numbers, number of patients waiting and the number of lost appointment slots including numbers of failures, cancellations and patients that did not attend (DNAd), often split according to procedure type. Data

were used by the endoscopy staff to monitor services and were also uploaded to the MES programme team on a monthly basis for external analysis for the final report published by the NHSMA.

For those sites that were unsuccessful in their application to participate in the MES programme, it was important that their enthusiasm was still rewarded and that help was given wherever possible, because these endoscopy units also had to meet the same targets as those participating in the MES programme. These sites were all offered advice and support, access to the Toolkit™ and NHSMA Improvement Guides if they wished to redesign their services independently. However, no funding was offered by the NHSMA and any data collected using the Toolkit™ was for local use only and was not required by the NHSMA.

Based on the data downloaded to the NHSMA, the MES programme was considered to be an outright success. The NHSMA published a report with numerous case studies illustrating the successful attainment of a number of targets [4]. However, it was possible that the report was bias in its findings, given the subjective source of the data on which the conclusions were made. To address this possibility, an independent, objective evaluation was funded by the National Institute of Health Research Service Delivery and Organisation to determine the long-term impact of the MES programme – the Evaluating New Innovations in Gastroenterology service delivery by the NHS Modernisation Agency (ENIGMA) study (SDO 46/2003). To achieve this, it would objectively compare the endoscopy services of a random selection of sites participating in the MES programme (intervention sites) with the endoscopy services in sites that were unsuccessful in their application to participate in the MES programme but who intended to redesign independently (control sites).

The ENIGMA study was a mixed methods study with patient quality of life (QoL) as a primary outcome measure [5]. Secondary outcome measures included interviews with patients and health professionals, a cost-effectiveness analysis, a GP survey, focus groups and routinely collected data. The ENIGMA study reported that there were no significant differences between the intervention and control sites and while improvements were made in the intervention sites, they were closely mirrored to improvements in the control sites. They concluded that the MES programme did not significantly improve NHS endoscopy services above and beyond what could have been achieved with just the intention to redesign. Papers specific to each component of the study are currently under development or being reviewed elsewhere.

This paper reports in detail on the ENIGMA study's innovative use of routinely collected, service-related endoscopy data similar to that collected by the Toolkit™ to evaluate the MES programme.

Methods

Hospital recruitment

Of the 26 intervention sites participating in the MES programme, the ENIGMA study randomly selected 10 by interval choice using an assigned random number after ranking according to bed size to ensure stratification by size. Of the 70 control sites that were unsuccessful in their application to participate in the MES programme, 27 had notified the NHSMA of their intention to redesign their endoscopy services independently. The ENIGMA study selected 10 of these 27 in the same way as described for the intervention sites above. Demographic details of each of the study sites can be found in the ENIGMA main report [5].

Data capture

Following the withdrawal of one intervention site, nine intervention sites and 10 control sites were asked to submit routinely collected service-related data defined as referral numbers, number of patients waiting more than 3 months (wait >3 months), total number of patients waiting (snapshot), number of lost appointment slots (lost slots) and activity. These five measures were collected retrospectively for eight calendar months: January, June and December 2003, April and November 2004, April and October 2005 and April 2006. The 2003 dates coincided with the start, middle and end of the MES programme while the other dates corresponded with waves of patient recruitment by the ENIGMA study and would allow this data to be included in the modelling of patient QoL scores. Data were collected at two time points – all 2003 and 2004 datasets were collected in April 2005 and all 2005 and 2006 datasets were collected in June 2006 – to reduce the quantity of data being requested at any one time in the hope that it would be more likely to be returned, especially because the data request was not incentivized in any way.

Data collection was confined to the three most commonly requested endoscopies, namely Upper Gastrointestinal Endoscopies (UGEs), Flexible Sigmoidoscopies (FS) and Colonoscopies. These data were collectively referred to as 'split procedures' data. From these three procedure types, a 'total procedures' dataset was also calculated.

It soon emerged that many endoscopy units did not routinely collect any service-related data, an issue already discussed elsewhere [6]. Consequently, the same data request was made to the corresponding Trust Information Services (TIS) Departments of all sites to achieve a complete dataset.

Data were excluded if they did not conform to the request made, either in format (percentages instead of actual numbers) or in the specification for it to be split according to procedure type. Data were also excluded if it were deemed to be Trust-wide data as opposed to the hospital-specific datasets specified. Data outliers were excluded if deemed to be improbable data entry errors by the Trust or endoscopy unit.

Data accuracy

Hospital Episode Statistics (HES) data has been used in many health services research studies in NHS trusts in England associated with gastrointestinal disorders [7–11] although there was only one reported case to date of its use in investigating NHS endoscopy services [12]. Given its wide application for the measurement of NHS management patterns, HES Activity datasets were considered to be the most appropriate and best available datasets against which to validate the Activity data submitted by each study site by simple comparison by calculating the difference and by statistical comparison using Wilcoxon Signed Rank tests.

Hospital Episode Statistics Activity data for each procedure type were secured from the Health Information Research Unit (HIRU) at Swansea University. All data were matched on OPCS-4 codes to ensure the study and HES datasets were as comparable as

possible. Both datasets were compared split according to procedure type, site type and calendar months collected. Statistical tests were done using SPSS version 13 (Lead Technologies Inc, Chicago, IL, USA). A P value of ≤ 0.05 was considered to be statistically significant.

Data analysis

Data were aggregated according to site type (intervention and control) to compare intervention and control groups at specific points in time and within each group over time. Data were aggregated into three time points (2003, 2004 and 2005/06) to improve the accuracy of the data in light of the problems with missing data, and to provide a more representative mean value of the group for that time period. Total procedures data and split procedures data for each measure were analysed using a two-way analysis of variance (ANOVA) using SPSS. Post hoc analyses for significant within-groups differences were investigated *a posteriori* using Friedman's tests.

Results

Data availability

Only eight of 19 endoscopy units submitted at least one of the five measures requested while all TIS contacts supplied at least one of the five measures. Table 1 indicates data availability in the final dataset. Data were limited at earlier and later time points, resulting in the necessity to aggregate data according to year to provide more accurate values.

Data accuracy

The only dataset that could be validated was the Activity data. Only sites with just one endoscopy unit within the Trust could be properly compared with the HES data as this would have contained Trust-based data that were always not split down to hospital level in the past. Only eight sites had just one endoscopy unit per Trust and as such, were the only sites eligible for comparison.

Using data split according to procedure type, site type and time, we found a significant difference between study Activity data and HES Activity data. Closer examination of the raw data identified grossly underestimated HES data for two intervention sites as a result of Trusts designating endoscopy procedures as outpatient procedures instead of day cases. Once these two sites were excluded from the analysis, there were no significant differences between study and HES Activity data.

Data analysis

When data were plotted for each site individually, the trends were highly variable over time for all five measures. The variation reduced for each measure when data were aggregated according to site type and time to produce mean values for the intervention group (see Table 2) and the control group for 2003, 2004 and 2005/06 (see Table 3).

Table 1 Source of each dataset used in this study

Site ID	Referral numbers	Number of patients waiting (Wait >3 months & Snapshot)	Lost slots	Activity
Intervention sites				
1	Endoscopy	Endoscopy	Endoscopy	Endoscopy
4	Joint	Joint	ND	Joint
6	Endoscopy	Endoscopy	Joint	Joint
7	Joint	Endoscopy	ND	TIS
8	TIS	TIS	ND	TIS
11	TIS	TIS	TIS	TIS
13	TIS	TIS	TIS	TIS
16	Joint	TIS	TIS	TIS
18	TIS	TIS	TIS	TIS
19	TIS	TIS	TIS	Endoscopy
Control sites				
2	Endoscopy	Endoscopy	Endoscopy	TIS
3	TIS	TIS	TIS	TIS
5	Endoscopy	TIS	Endoscopy	Endoscopy
9	TIS	TIS	TIS	TIS
10	ND	Endoscopy	ND	TIS
12	Endoscopy	Endoscopy	ND	Endoscopy
14	TIS	TIS	TIS	TIS
15	TIS	TIS	TIS	TIS
17	Endoscopy	Endoscopy	Endoscopy	Endoscopy
20	Joint	Joint	ND	TIS

Endoscopy unit, TIS contact or joint effort from both.
ND, No data collected from any source; TIS, Trust Information Services.

Table 2 Mean ± SD for the intervention group using total procedures and split procedures data

Outcome measure	Procedure type	Intervention group mean ± SD (*n*) 2003	2004	2005/06
Referral numbers	FS	75 ± 51 (25)	75 ± 47 (18)	81 ± 55 (24)
	Colonoscopy	128 ± 107 (25)	134 ± 94 (18)	125 ± 70 (24)
	UGEs	259 ± 107 (25)	237 ± 97 (18)	218 ± 93 (24)
	Total procedures	462 ± 202 (25)	447 ± 160 (18)	423 ± 147 (24)
Wait >3 months	FS	19 ± 25 (15)	15 ± 13 (12)	11 ± 10 (15)
	Colonoscopy	65 ± 79 (15)	52 ± 55 (12)	60 ± 65 (15)
	UGEs	50 ± 81 (15)	49 ± 65 (12)	54 ± 89 (15)
	Total procedures	134 ± 154 (15)	116 ± 118 (12)	124 ± 145 (15)
Snapshot	FS	138 ± 95 (12)	131 ± 124 (8)	86 ± 88 (10)
	Colonoscopy	219 ± 152 (12)	191 ± 97 (8)	173 ± 86 (10)
	UGEs	328 ± 167 (12)	266 ± 75 (8)	303 ± 212 (10)
	Total procedures	685 ± 270 (12)	589 ± 130 (8)	561 ± 266 (10)
Lost slots	Total procedures	116 ± 97 (15)	127 ± 82 (10)	109 ± 113 (15)
Activity	FS	75 ± 58 (27)	74 ± 60 (18)	75 ± 48 (22)
	Colonoscopy	138 ± 107 (27)	155 ± 114 (18)	131 ± 155 (22)
	UGEs	256 ± 93 (27)	255 ± 95 (18)	235 ± 71 (22)
	Total procedures	469 ± 167 (27)	484 ± 159 (18)	441 ± 161 (22)

Sample numbers are denoted in brackets.
FS, Flexible Sigmoidoscopies; UGEs, Upper Gastrointestinal Endoscopies.

Table 3 Mean ± SD for the control group using total procedures and split procedures data

Outcome measure	Procedure type	Control group mean ± SD (*n*) 2003	2004	2005/06
Referral numbers	FS	38 ± 46 (9)	86 ± 66 (13)	94 ± 58 (16)
	Colonoscopy	135 ± 37 (9)	122 ± 47 (13)	132 ± 73 (16)
	UGEs	294 ± 95 (9)	244 ± 107 (13)	250 ± 104 (16)
	Total procedures	467 ± 147 (9)	453 ± 186 (13)	476 ± 187 (16)
Wait >3 months	FS	8 ± 8 (7)	7 ± 11 (7)	10 ± 16 (10)
	Colonoscopy	127 ± 137 (7)	120 ± 175 (7)	76 ± 213 (10)
	UGEs	32 ± 26 (7)	70 ± 107 (7)	41 ± 122 (10)
	Total procedures	167 ± 163 (7)	197 ± 291 (7)	127 ± 340 (10)
Snapshot	FS	38 ± 48 (8)	150 ± 243 (10)	94 ± 39 (10)
	Colonoscopy	573 ± 333 (8)	388 ± 355 (10)	631 ± 449 (10)
	UGEs	447 ± 299 (8)	511 ± 284 (10)	665 ± 418 (10)
	Total procedures	1058 ± 594 (8)	1049 ± 588 (10)	1390 ± 844 (10)
Lost slots	Total procedures	94 ± 67 (7)	66 ± 48 (8)	75 ± 49 (11)
Activity	FS	38 ± 29 (17)	64 ± 49 (16)	72 ± 57 (19)
	Colonoscopy	78 ± 30 (17)	90 ± 44 (16)	110 ± 66 (19)
	UGEs	237 ± 107 (17)	236 ± 94 (16)	220 ± 70 (19)
	Total procedures	352 ± 126 (17)	390 ± 142 (16)	401 ± 143 (19)

Sample numbers are denoted in brackets.
FS, Flexible Sigmoidoscopies; UGEs, Upper Gastrointestinal Endoscopies.

When examining the actual change in mean data from 2003 to 2005/06 according to site type, we found that the Referral numbers for the intervention group decreased while the control group showed a slight increase. The Wait >3 months in the intervention group decreased slightly but showed a more pronounced decrease in the control group. The intervention group Snapshot decreased but increased in the control group. The intervention and control group Lost slots both decreased slightly. Finally, the Activity data for the intervention group decreased while the control group activity increased.

The two-way ANOVA using Total procedures data revealed no statistically significant between-groups or within-groups effects for any of the five measures analysed (see Table 4). This indicated that any changes in data in the intervention group over time were not statistically significant and were mirrored by equivalent changes in the data of the control group. Activity showed a significant interaction effect ($F_{2,26} = 3.594$, $P = 0.042$), indicating that there was a significant, non-parallel difference in the direction of change in the mean Activity data over time in the intervention and control groups.

When using split procedures data, there were no significant between-groups effects for any of the five outcome measures (see Table 4). The only significant within-group differences over time were for UGE-specific Referral numbers ($F_{1,11} = 5.15$, $P = 0.03$)

Table 4 Two-way ANOVA using total procedures and split procedures data

Outcome measure	Procedure type	Within-subject effects (*F* ratio, sig.)	Between-subject effects (*F* ratio, sig.)	Interaction effects (*F* ratio, sig.)
Referral numbers	FS	2.12, 0.169	0.059, 0.813	0.94, 0.365
	Colonoscopy	0.347, 0.58	0.201, 0.663	0.733, 0.417
	UGEs	5.151, **0.03**	0.646, 0.439	0.284, 0.67
	Total procedures	0.28, 0.64	0.586, 0.46	0.317, 0.617
Wait >3 months	FS	0.016, 0.984	0.246, 0.638	2.296, 0.143
	Colonoscopy	0.965, 0.367	1.004, 0.355	1.247, 0.308
	UGEs	0.992, 0.362	0.226, 0.651	0.896, 0.385
	Total procedures	0.994, 0.36	0.594, 0.47	1.29, 0.3
Snapshot	FS	0.313, 0.737	0.826, 0.399	4.428, **0.036**
	Colonoscopy	0.931, 0.421	4.351, 0.082	3.06, 0.084
	UGEs	0.812, 0.467	3.23, 0.122	1.235, 0.325
	Total procedures	0.733, 0.435	3.757, 0.101	2.461, 0.163
Lost slots	Total procedures	0.313, 0.737	0.469, 0.519	0.965, 0.409
Activity	FS	0.019, 0.981	0.899, 0.36	3.14, 0.06
	Colonoscopy	2.928, 0.103	0.255, 0.622	2.259, 0.151
	UGEs	5.249, **0.012**	0.586, 0.458	1.789, 0.187
	Total procedures	0.348, 0.71	1.106, 0.312	3.594, **0.042**

All significant values are highlighted in bold.
FS, Flexible Sigmoidoscopies; UGEs, Upper Gastrointestinal Endoscopies.

and UGE Activity ($F_{1,13} = 5.25$, $P = 0.012$), indicating that those two measures changed significantly over time within the intervention and control groups. Further analysis split according to site type using the Friedman's test revealed a significant difference over time in the control group UGE-specific Referral numbers ($n = 4$, d.f. = 2, $\chi^2 = 6$, $P = 0.05$), but not for the intervention group UGE-specific Referral numbers ($n = 9$, d.f. = 2, $\chi^2 = 4.67$, $P = 0.097$). Friedman's tests also revealed significant differences in the UGE-specific Activity data for the intervention group ($n = 8$, d.f. = 2, $\chi^2 = 7$, $P = 0.03$) but not for the control group ($n = 7$, d.f. = 2, $\chi^2 = 2$, $P = 0.368$).

The only significant interaction effect for split procedures data was for the FS-specific Snapshot data ($F_{1,6} = 4.43$, $P = 0.036$), indicating that there was a significant, non-parallel difference in the direction of change of this data over time in the intervention and control groups.

Discussion

This study found that there were no significant differences between intervention and control groups for Referral numbers, Wait >3 months, Snapshot, Lost slots and Activity. The only significant within-groups differences were in the control group for UGE-specific Referral numbers and within the intervention group for UGE-specific Activity. There were significant interaction effects for total procedures activity data and for the FS-specific Snapshot data.

Unfortunately, while there is a plethora of literature in the field of general health services evaluation, no other studies have been done in this specific field to allow a comparative discussion. The only comparable studies are the other components of the ENIGMA study, all of which are complementary in their findings to this study.

This study has an innovative design and strict inclusion and exclusion criteria to ensure the final datasets were as accurate as possible and made up of comparable datasets and to minimize potential methodological and outcome biases and confounding. However, data availability was problematic for this study, although the inclusion of Trust data and the aggregation of data for statistical analyses should have reduced its impact.

The decision to only collect eight time points-worth of data was based on the potential implications of this fairly resource-intensive data request. There was no funding for this data request and because many sites did not collect data, it would have increased their workload and possibly caused them to withdraw from the ENIGMA study. Ideally, a time series analysis would have been done using data collected for each calendar month or on a bimonthly basis to better illustrate data trends. The compromise was to select data for only eight time periods that tied in with the patient recruitment so that the data could be used in the analysis presented here, as well as in a modelling exercise with the patient QoL scores.

The ENIGMA study evaluated a complex intervention – the MES programme – and to do this effectively and realistically, a range of hospital sizes were necessary. The data in this study were not standardized prior to analysis but we believe that by analysing the mean data for the intervention and control groups, we reduce confounding according to hospital bed size. Also, the selection process whereby sites were ranked according to bed size would have ensured approximately equal numbers of each size of hospital in both study groups.

The decrease in Referral numbers in the intervention group may have been due to the MES programme advocating the management of referrals using validation procedures and new referral pathways. Both site types showed decreases in the number of patients waiting more than 3 months over time, as was expected since it was an NHS Cancer Plan target [2]. The decrease in the Snapshot data in the intervention group may be attributed to waiting list validation and pooling, while the concurrent increase in the control group

may have been due a reclassification of patients to meet local and national cancer targets. The control group were better at reducing their already lower lost slots data over time than the intervention group, although there is no obvious reason why this would be the case. Perhaps the high profile nature of DNAs and cancellations in the NHS in general made this an obvious priority in the absence of Toolkit™ data with which to properly identify problem areas. Finally, the decrease in activity over time in the intervention group may be due to more endoscopic procedures being done as outpatient procedures, a measure not recorded by this study, while the control group increase in Activity may have been due to the decrease in the number of Lost slots meaning more procedures being counted as completed.

There may be any number of reasons why the endoscopy services of the intervention sites did not significantly improve over time, including a lack of ongoing, high-quality data collection, a lack of ongoing external funding and support from the NHSMA, the constant internal and external pressures to meet new targets and the lack of sustainability of the MES-initiated reforms. It is also possible that, given the strict nature and intensity of the MES programme, many sites were experiencing 'reform exhaustion' and felt unwilling to sustain redesign efforts or to proceed with either planning or implementing any further modernization programmes.

It was not surprising that the control sites did not significantly improve their services, because they had limited NHSMA support and no MES funding during 2003 but were expected to address many of the same external and internal pressures. Control sites fully intended to redesign and had access to web-based NHSMA redesign literature and the Toolkit™ so it was reasonable to anticipate some improvements in their services, but at a later date than the intervention sites and possibly at a more fundamental level. However, the MES programme may have acted as a catalyst in these sites following their in-house evaluation of their services as part of their application for the MES programme.

It is important to evaluate NHS services to ensure that any process is running optimally and to guarantee that there is no alternative way of doing things that would be more efficient. An effective way to do this is to look at demand, capacity and activity and determine how well-matched they are [3]. This was the primary function of the Toolkit™ during the MES programme and allowed both sites and the NHSMA to analyse their services prior to redesigning them and to measure the impact of any redesign programmes on the service in the short- and long-term. The evaluation method used in this study was based on the same principles used by the Toolkit™, namely the analysis of routinely collected, service-related data to examine endoscopy services and changes thereof over time.

This study found that only three intervention sites routinely collected at least one of the measures requested, an issue discussed elsewhere [6]. There is a clear message here for all externally led modernization agendas – no matter how good the concept, there is no guarantee that it will remain in use once it becomes voluntary. This means that future modernization programmes will need to consider not only how they encourage NHS services to redesign and improve their services, but how they will sustain modernization and the prolonged use of any ideas or tools after the programmes close. In the case of the Toolkit™, its complexity and rigorousness actually added to the workload of the endoscopy staff and so, was never likely to be sustained long-term for that reason. New modernization concepts need to be time- or resource-savers, adapted for more practical use and easily embeddable into everyday use in the service so that it takes more effort to withdraw them from use than to keep them.

This study has identified some important findings relating to the impact of the MES programme and the availability, or lack therefore, of service-related data. However, it is important that readers understand that all findings and conclusions drawn from this study have been done using limited datasets and using a study design that was focussed on providing adequate statistical power for the patient QoL measures that were the primary outcome measure for the ENIGMA study, and not the analysis of service-related data.

We conclude that the MES programme did not significantly improve NHS endoscopy services over and above what could have been achieved with only the intention to redesign. Also, data were not routinely collected in these study sites, illustrating the inability of many NHS endoscopy services to proactively measure and evaluate their services. The NHS needs to adopt an ethos of high-quality data collection and analysis throughout the organization, across primary, secondary and tertiary care boundaries for the whole of the UK if it hopes to make and sustain any measurable improvements in its delivery of care to patients.

Acknowledgements

The ENIGMA study was funded by the National Institute for Health Research, Service Delivery and Organisation Programme (SDO 46/2003).

References

1. Department of Health (2000) The NHS Plan: A Plan for Investment. A Plan for Reform. London: Department of Health.
2. Department of Health (2006) Quality and value: quick wins. *Health Service Journal Supplement*, 6 April 2006.
3. NHS Modernisation Agency (2003) Modernising Endoscopy Services – A Practical Guide to Redesign. Leicester: NHS Modernisation Agency.
4. NHS Modernisation Agency (2004) National Endoscopy Programme: Report from the Second Wave. Leicester: Modernisation Agency.
5. Williams, J. G., Cheung, W. Y., Cohen, D., Hutchings, H., Jerzembek, G., Rapport, F., Russell, I. T. & Seagrove, A. C. (2008) Evaluating Innovations in the Delivery and Organisation of Gastroenterology Services Initiated Directly or Indirectly by the Modernising Endoscopy Services Programme of the NHS Modernisation Agency. London: National Institute for Health Research Service Delivery and Organisation Programme.
6. Thorne, K., Hutchings, H. A. & Elwyn, G. (2008) Unmeasured improvement work: the lack of routinely collected, service-related data in NHS endoscopy units in England involved in 'modernisation'. *BMC Health Services Research*, 8, 20.
7. Pollock, A. M. & Vickers, N. (1998) Trends in colorectal cancer care in southern England, 1989–1993: using HES data to inform cancer services reviews. *Journal of Epidemiology and Community Health*, 52 (7), 433–438.
8. Pollock, A. M. & Vickers, N. (1998) Deprivation and emergency admissions for cancers of colorectum, lung, and breast in south east England: ecological study. *British Medical Journal*, 317 (7153), 245–252.
9. Kang, J. Y., Hoare, J., Tinto, A., Subramanian, S., Ellis, C., Majeed, A., Melville, D. & Maxwell, J. D. (2003) Diverticular disease of the colon

– on the rise: a study of hospital admissions in England between 1989/1990 and 1999/2000. *Alimentary Pharmacology & Therapeutics*, 17 (9), 1189–1195.
10. Parry, J., Jolly, K., Rouse, A. & Wilson, R. (2004) Re-organizing services for the management of upper gastrointestinal cancers: patterns of care and problems with change. *Public Health*, 118 (5), 360–369.
11. Al-Sarira, A. A., David, G., Willmott, S., Slavin, J. P., Deakin, M. & Corless, D. J. (2007) Oesophagectomy practice and outcomes in England. *British Journal of Surgery*, 94 (5), 585–591.
12. Williams, J. G. & Mann, R. Y. (2002) Hospital episode statistics: time for clinicians to get involved? *Clinical Medicine*, 2 (1), 34–37.

