JOURNAL OF INNOVATION IN HEALTH INFORMATICS

Research article

How do clinical information systems affect the cognitive demands of general practitioners? Usability study with a focus on cognitive workload

Ferran Ariza
Independent General Practitioner, London, UK

Dipak Kalra
Centre for Health Informatics and Multiprofessional Education, University College London, London, UK
The EuroRec Institute, Gent, Belgium

Henry W W Potts
Centre for Health Informatics and Multiprofessional Education, University College London, London, UK

Cite this article: Ariza F, Kalra D, Potts HWW. How do clinical information systems affect the cognitive demands of general practitioners? Usability study with a focus on cognitive workload. *J Innov Health Inform*. 2015;22(4):379-390.

http://dx.doi.org/10.14236/jhi.v22i4.85

Copyright © 2015 The Author(s). Published by BCS, The Chartered Institute for IT under Creative Commons license http://creativecommons.org/licenses/by/4.0/

Author address for correspondence:
Ferran Ariza
Centre for Health Informatics and Multiprofessional Education,
University College London, 222 Euston Road,
London NW1 2DA, UK
E-mail: feariza@doctors.org.uk

Accepted September 2015

ABSTRACT

Background Clinical information systems in the National Health Service do not need to conform to any explicit usability requirements. Poor usability can increase the mental workload experienced by clinicians and cause fatigue, increase error rates and impact the overall patient safety. Mental workload can be used as a measure of usability.

Objective To assess the subjective cognitive workload experienced by general practitioners (GPs) with their systems. To raise awareness of the importance of usability in system design among users, designers, developers and policymakers.

Methods We used a modified version of the NASA Task Load Index, adapted for web. We developed a set of common clinical scenarios and computer tasks on an online survey. We emailed the study link to 199 clinical commissioning groups and 1,646 GP practices in England.

Results Sixty-seven responders completed the survey. The respondents had spent an average of 17 years in general practice, had experience of using a mean of 1.5 GP computer systems and had used their current system for a mean time of 6.7 years. The mental workload score was not different among systems. There were significant differences among the task scores, but these differences were not specific to particular systems. The overall score and task scores were related to the length of experience with their present system.

Conclusion Four tasks imposed a higher mental workload on GPs: 'repeat prescribing', 'find episode', 'drug management' and 'overview records'. Further usability studies on GP systems should focus on these tasks. Users, policymakers, designers and developers should remain aware of the importance of usability in system design.

Keywords: cognitive science, computerized medical record systems, general practice, human engineering, medical informatics, primary health care, user–computer interface.

What does this study add?

- Current GP systems in England do not need to conform to explicit usability requirements. Poor usability can increase the mental workload of clinicians and lead to errors.
- Some clinical computer tasks incur more cognitive workload than others and should be considered carefully during the design of a system.
- GPs did not report overall very high levels of subjective cognitive workload when undertaking common clinical tasks with their systems.
- Further usability studies on GP systems should focus on the tasks incurring higher cognitive workload.
- Users, policymakers, and designers and developers should remain aware of the importance of usability in system design.

INTRODUCTION

Modern computer information systems have a proven record in supporting information management tasks in healthcare.[1,2] However, clinical information systems lacking usable designs may increase the mental workload imposed on their users and negatively impact on patient safety.[3–5] Today, with the nearly universal computerization of general practice in the UK,[6] general practitioners (GPs) are faced with the challenge of working with computer systems that could be potentially adding to their cognitive workload.[7]

GP Systems of Choice (GPSoC) in England (Table 1) do not have to conform to any specified usability requirements, although there have been initiatives to promote usability supported by the National Health Service (NHS), including the Microsoft Health Common User Interface.[8] Usability has been defined as 'the capacity of a system to allow users to carry out their tasks safely, effectively, efficiently and enjoyably'.[9] It largely relates to the quality of the information design and system navigability, and has a lot to do with user interfaces.[10,11] Usability is a very important aspect in system design[12] and may impact user satisfaction, user fatigue, clinical productivity, error rates, and overall patient safety.

Table 1. GP systems and suppliers in the NHS

GPSoC framework supplier: GPSoC compliant system/s
CSC Computer Sciences Limited: SystmOne
EMIS: LV, PCS and WEB
Advanced Health and Care: Crosscare
InPractice: Vision 3
iSOFT: Premiere and Synergy
Microtest: Evolution and Practice Manager II

Clinical care is a complex and mentally demanding activity. It usually involves multitasking, where communication skills[13] are combined with pattern recognition, clinical reasoning and problem solving skills.[14] In primary care, the high levels of uncertainty,[15] the variety of clinical presentations seen[16] and need to deal with vast amounts of information routinely received from other care settings[17] add to this complexity. All of these contribute to the mental workload experienced in general practice.[18]

Mental workload has been defined as the mental effort involved in performing any given task.[19] Excessive simultaneous tasks or very demanding individual tasks may cause cognitive overload.[20] In healthcare, multitasking may lead to errors.[21] Information systems should help users complete tasks with ease. Their user interfaces should be carefully designed to ensure that the way information is displayed does not negatively affect the cognitive workload experienced by healthcare professionals.[22] Improved user interfaces should reduce medical errors and improve patient care.[23]

Scottish GPs have expressed reservations about the usability of their computer systems.[24] Usability issues in healthcare have been noted elsewhere. Examples cited are fixed sequences of steps and tasks, poorly supported documentation and retrieval of information,[25] problems relating to the efficiency of use, intuitiveness, and poor fit for multiprofessional needs,[26] and poor integration into workflow.[27] In healthcare information technology, the main challenge today appears to lie on the design of convenient, efficient and acceptable interfaces.[28] Failure of many clinical information systems has been attributed to the lack of human-centered design approaches.[29,30] Problems with usability have been linked to patient safety elsewhere.[31–33] In the UK, poor usability has also been identified as a safety issue.[34,35]

Guidelines for the evaluation of the usability of electronic record systems in primary care exist.[36] Among several existing validated subjective mental workload tools,[37] NASA-TLX[38] has been commonly used in health care. This is a multidimensional scale consisting of two parts: 1) six 20-point visual analogue sub-scales that measure contributing factors to the overall cognitive load (measures of workload): mental demand, physical demand, temporal demand, performance, effort and frustration; and 2) a set of pairings combining the different factors to be rated according to relevance, in order to weight the

subjective importance given to the difference factors (sources of workload). This has been used in a variety of health information and communication technology evaluations such as nursing intensive care,[39] anesthetic work,[19] surgical training and surgery,[40,41] and electronic white boards.[42]

This study looks to identify where potential usability issues with current existing GP systems in the NHS in England may exist, by exploring the subjective cognitive workload experienced by GPs when they carry out common clinical tasks with their computer systems. The study hopes to raise awareness of the influence of usability on safety among system users, developers and policy makers.

METHODS

The study presented participants with 10 stories each describing a clinical scenario with a related computer task. GPs rated the subjective mental workload experienced with their system for that task, using the tool described below.

Setting

The study was carried out in the primary care setting in England.

Population

Eligible participants were GPs who at the time of the study: 1) were working or had worked within the past 6 months, for the NHS or the Ministry of Defence in England and 2) were using or had used one or more of the currently existing GP systems in a paperless or paper-light practice/setting. Participation in the study was voluntary. Data were anonymously collected through an online survey.

Materials

NASA Task Load Index (NASA-TLX)[43,44] is a tool designed to undertake subjective workload assessments and is free to use. We used a modified version (known as Raw TLX[38]), consisting of six 20-point visual analog sub-scales that measure six contributing factors to the overall cognitive load (measures of workload): mental demand, physical demand, temporal demand, performance, effort and frustration (Figure 1). For this study, we dropped off the physical demand scale that we judged irrelevant.

We transformed the tool for a web-based form using Google Forms. The survey contained instructions for the study, demographic questions, clinical scenarios and tasks requested and rating scales with descriptions of the measure. It also included a space for free comments after each task (Figure 2).

In order to enable participants to complete the study at a time of their convenience, we presented them with a set of common clinical scenarios and tasks and asked them to recall (rather than perform) how they use their systems to complete these tasks. This approach ensured that time involvement was minimized and prevented interference with work.

We collected a number of demographic variables (time in general practice, time using the system and the number of systems used) in order to identify if any of these factors have an impact on cognitive load.

Measures

Each questionnaire returned 50 scores in total: five scores corresponding to five sub-scales ratings for each of the 10 tasks. We refer to the average of these 50 scores as the 'survey score'. For each task, we computed a single average score, which we refer to as the 'task score'. Scores are given

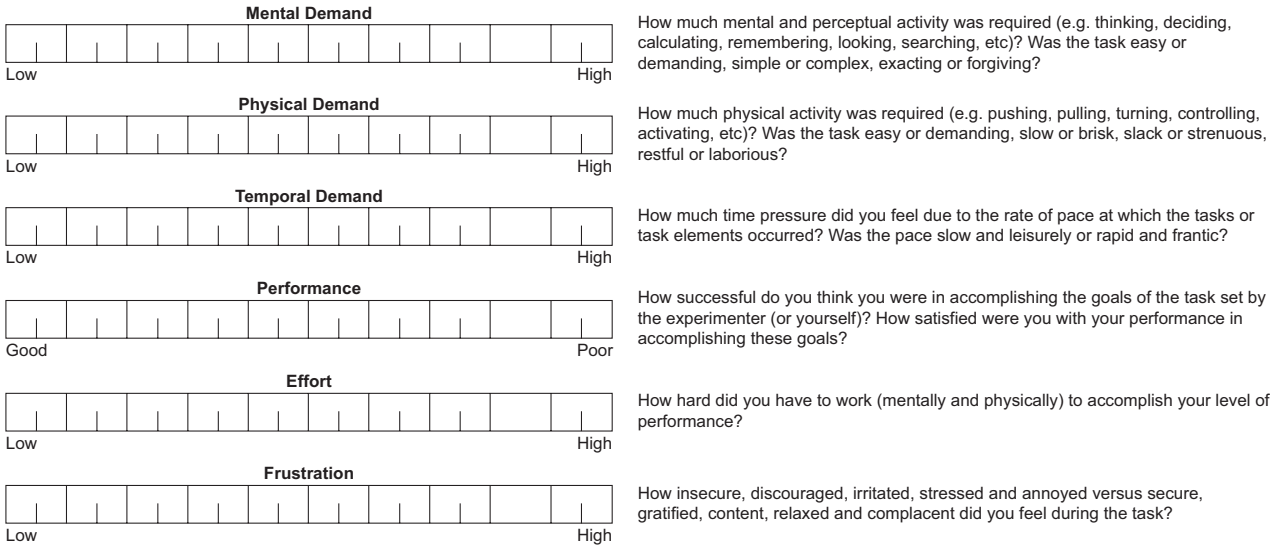

Figure 1. NASA-TLX dimensions and rating scales

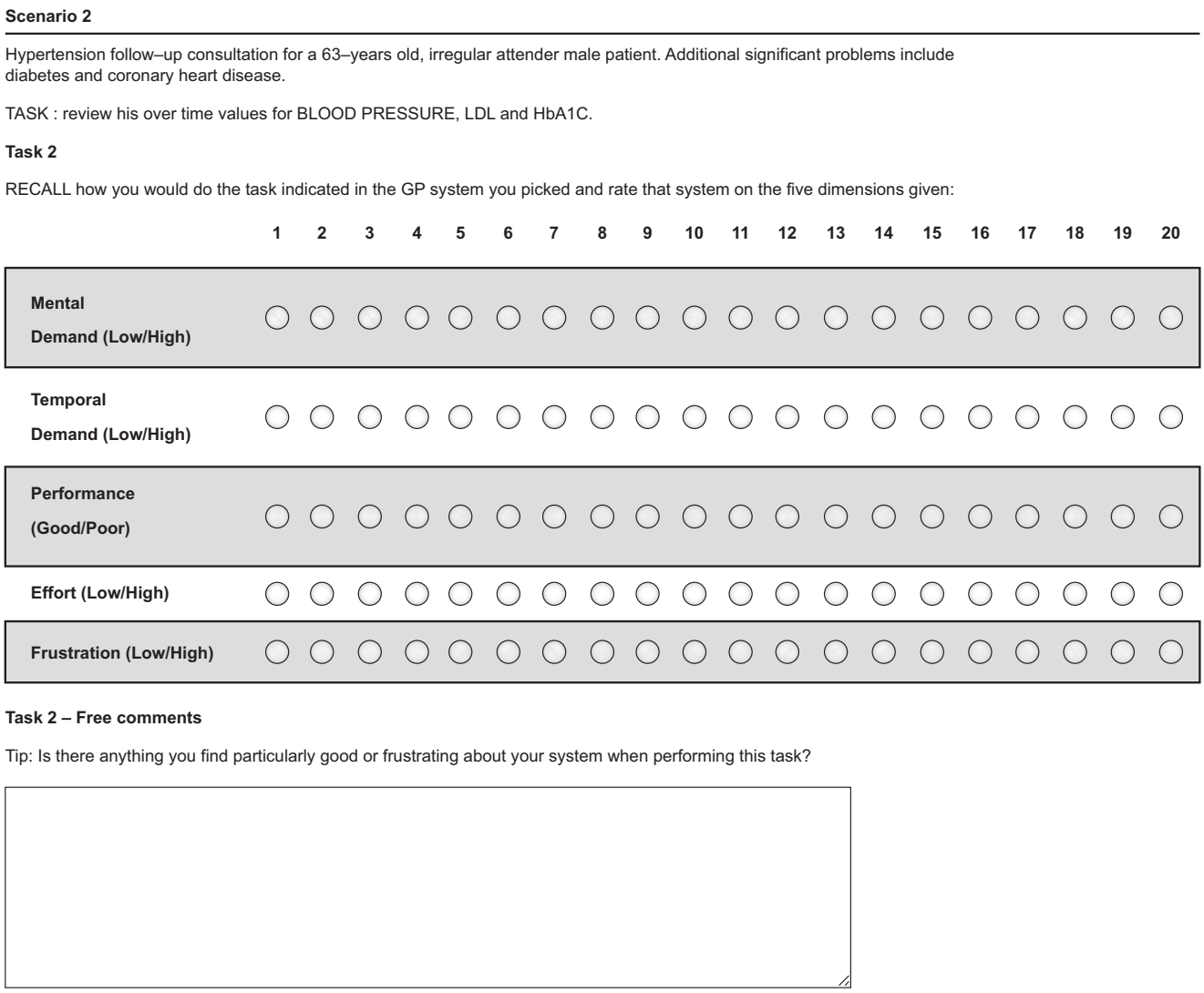

Figure 2. Web version of a scenario and NASA-TLX rating scales

in a scale 0–100. Additional measures were the number of years in general practice, the number of years the system had been used and the number of other systems used.

Selection, design and validation of clinical scenarios and tasks

We made a selection of tasks using Laerum and Faxvaag's task list,[45] with some additions judged to be relevant to GP clinical information systems. The tasks selected belong to the generic following activity groups: retrieving information, entering information, prescribing and managing prescriptions and dealing with laboratory results. Following Laerum and Faxvaag,[45] we use 'task' here as a high-level description. Each task involves multiple actions, each of which contributes to the overall cognitive workload. These will involve the user interacting with the information system using keyboard and mouse (or equivalent) and making various decisions.

We adapted the clinical scenarios from real cases, taking into account the available data on the most frequent presenting complaint groups and most frequently prescribed drugs.[46] We ensured the final validity of the scenarios by asking participants about the representatives of the cases during a pilot test. The final list of clinical scenarios and tasks is available in Table 2.

Table 2. Clinical scenarios and tasks

Number	Task description	Brief descriptor
1	Home visit request for an 82-year old lady who is 'confused'. Task: in preparation for the visit, please check: current/active problems, past relevant/significant problems, active prescriptions (acute and repeat), recent consultation/s, recent lab results.	Overview records
2	Hypertension follow-up consultation for a 63-year-old, irregular attender male patient. Additional significant problems include diabetes and coronary heart disease. Task: review over time values for blood pressure, LDL and HbA1C.	Values over time

3	22-year-old female patient tells you that she has had depression in the past and she thinks this is coming back. She was given an antidepressant before but she doesn't think it helped much. Task: find previous existing entry for the depressive episode and medication prescribed.	Find episode
4	52-year-old female with a 24-hours history of uncomplicated gastroenteritis. Task: record the problem using a clinical code for 'Gastroenteritis'. Enter the history, examination and management details using free text.	Record unstructured
5	42-year-old patient, smoker of 20 cig/ day, with acute respiratory symptoms for 3 days. Temperature 37.4 C, BP 142/91, pulse 92 bpm. Task: record the problem using a clinical code for 'lower resp tract infection'. Enter the history, physical examination and management details using free text. Record his temperature, blood pressure, pulse and smoking status using structured data entry.	Record structured
6	Issue the following acute prescriptions for a 75-year-old moderate COPD patient with an acute exacerbation: Amoxicillin 500 mg capsules (Supply 21 capsules) Take one capsule three times a day for 7 days Prednisolone 5mg gastro-resistant tablets (Supply 42 tablets) Take six tablets each morning (as a single dose) for 7 days Salbutamol 100 micrograms/actuation inhaler CFC free (Supply 1 200 dose inhaler) Inhale one to two puffs up to four times a day, when required to relieve breathlessness.	Acute prescribing
7	62-year-old newly registered patient with a history of diabetes, hypertension and TIA needs his medications. Task: Set up these medications for regular repeat, and issue today's prescriptions (all needed): Amlodipine 10mg tablets (supply 56 tablets) Take one tablet once a day Authorize three issues Aspirin 75 mg dispersible tablets (supply 56 tablets) Take one tablet once a day Authorize three issues Atorvastatin 80 mg tablets (supply 56 tablets) Take one tablet daily Authorize three issues Lisinopril 10 mg tablets (Supply 56 tablets) Take one tablet once a day Authorize three issues Metformin hydrocholoride 500 mg tablets (Supply 224 tablets) Take two twice a day Authorize three issues Medication review in 6 months.	Repeat prescribing
8	Hypertension follow-up consultation with a 47-years male patient. Four weeks ago you added amlodipine and changed his simvastatin to pravastatin. Blood pressure today 141/89. LFTs and U&Es normal. He reports no side effects to the new medications. Task: amend repeat prescriptions as outlined below. Cancel the following: Simvastatin 40 mg tablets Take one tablet at night. Supply 28 tablets. Omeprazole 20 mg gastro-resistant capsules Take one capsule once a day. Supply 28 capsules. (Prescribed for indigestion; no longer using it) Salbutamol 100 micrograms/actuation inhaler CFC free (Supply 1 200 dose inhaler) Inhale one to two puffs up to four times a day (Prescribed it due to hay fever related wheezing ; no longer needed)	Drugs management

continued

	Add the following: Amlodipine 5 mg tablets (Supply 56 tablets) Take one tablet once a day Pravastatin 80 mg tablets (Supply 56 tablets) Take one tablet daily Authorize three issues Re-authorize the following (issue limit reached): Lisinopril 10 mg tablets Take one tablet once a day. Supply 28 tablets. Authorize three issues medication review in 6 months.	
9	A 45-year-old male patient comes today to find out the results of a recent blood test. Task: look up the results on the system.	View labs
10	You are reviewing your system labs inbox. One report shows: normal FBC, glucose and U&E; cholesterol 5.9 mmol/L and ALT 54 U/L. Task: add a comment advising the patient to come in to discuss the results. Mark the report as seen and file/archive as necessary.	Action labs

Survey design

An initial version of the survey was piloted by eight GPs. They confirmed the correct functioning of the online survey and appropriateness of the clinical scenarios. They made comments on the understanding of the scales and the scoring system, the clarity and level of detail of the clinical scenarios and the general design of the survey. The changes introduced resulted in a survey with 10 simplified clinical scenarios and three demographic questions for 'time in general practice', 'time using the system' and 'number of systems used'.

Sampling method and participants' recruitment

The target population included an estimated 40,000 GPs in England.[47] Due to funding limitations, we used existing electronic communication channels to try reaching the largest number of GPs possible and maximize responses rates. During late May and June 2013, we submitted an electronic request to 199 clinical commissioning groups, who are responsible for organising the delivery of NHS services in England, requesting assistance with the survey distribution. We also emailed directly 1,646 GP practices with an invitation to participate in the study.

Methods of analysis

We used SPSS for the statistical analysis. We looked at the correlation between all 50 variables obtained for each questionnaire, followed by dimension reduction by factor analysis. We computed 12 additional scores: 10 task scores, an average score for all tasks (referred to as survey score), and a score for the number of additional systems used.

We obtained means with confidence intervals for the survey scores and the tasks scores. We looked for statistical differences between the tasks and the systems using ANOVA for repeated measures. We also looked at the influence of the variables, 'time in general practice', 'time using system', and 'number of other systems used'.

RESULTS

Sixty-seven respondents completed the online survey. A precise estimate of the responses rate was not possible because we did not know the number of GPs who received the invitation. However, a gross estimate could be between 0.5 and 1%, based on a possible number of GPs receiving the survey of 5,000–10,000.

The distribution of systems used by participants was 55.2% for all combined EMIS systems (25.4% LV, 10.4% PCS and 19.4% Web), 29.9% for SystmOne, 9.0% for INPS Vision, 3.0 % for iSoft Synergy, 1.5% for Microtest Evolution and 1.5% for other systems not in the GPSoC approved list. The mean time the system had been used was 6.7 years, the average time in general practice 17.8 years and the average number of other systems used 1.5 (Table 3).

According to a report from 2011,[48] the market share of GP systems in England was 55% for EMIS, 19% for INPS Vision, 17% for TTP SytstmOne, 7% iSoft and 2% Microtest. More recent data suggested an EMIS market share of 54.8% and iSoft share of 5.6%;[49] also, TTP SystmOne was set to become the second biggest supplier.[50] Based on these details, we projected a possible current market distribution of the systems as follows: 54.8% for EMIS, 19.6% for TTP SytstmOne, 18% for INPS Vision, 5.6% iSoft and 2% Microtest. According to these figures, the system distribution of our sample was not significantly different from the population distribution, $\chi^2(4) = 7.64$, $p > 0.10$.

Although we did not have population data on the average time for GPs in general practice, the average time the system had been used and the average number of other systems used, a qualitative study in Scotland investigating the views of GPs on their medical records (n = 25 GPs) reported an average time in general practice of 16.5 years,[24] which is similar to the number found in our study.

We identified a potential problem with the 'performance' scale, which appeared to have been marked in the wrong direction in a number of cases. Some participants also indicated this in the free-text section. The NASA-TLX user guide document highlights a possible confusion with this scale. This has been reported in other studies.[23] We carried out a correlation analysis that also showed a problem with the scale for performance. Therefore, we dropped out this scale for the analysis. Otherwise, this showed that all the scores were highly correlated.

A dimension reduction by factor analysis showed one factor accounting for around 67% of the variance, where all the scales where highly and equally correlated within the

factor. Other studies have previously reported that scales are often significantly correlated with each other.[38] Since all the scales and tasks correlated well, we created a single aggregated score with the average of all 40 remaining ratings (after dropping the performance scale). We also computed a total score for each task, from averaging the scores of the four scales.

The overall cognitive workload score was 28.7 [23.3–34.0, 95% confidence interval (CI)]. This score was not significantly different among systems (F (4, 58) = 0.3, p = 0.88) (Microtest and iSoft were excluded due to small numbers) (Figure 3).

A repeated measures ANOVA of the total scores for each task revealed statistically significant differences between them (F (9, 58) = 6.1, p = 0.001). The data in Table 4 and the graph in Figure 4 show that the tasks 'overview records', 'find episode', 'repeat prescribing' and 'drugs management' scored significantly higher than the tasks 'record unstructured', 'view labs' and 'action results'. The tasks 'find episode' and 'repeat prescribing' also scored significantly higher than the task 'acute prescribing'. Finally, 'the values over time' and 'the record structured' did not score significantly different from any of the other tasks.

The tasks, however, were not significantly different among the different systems in a repeated measures ANOVA; test for interaction between system and task: (F (36, 189.1) = 0.7, p = 0.9). The difficulty of the tasks was not related to the time the GPs had been in general practice when that was included as a covariate in a repeated measures ANOVA (F (1, 55) = 0.3, p =0.58). The same was true of the number of other systems used (F (1, 55) = 0.03, p =0.87), but difficulty was related to the time the system had been used for (F (1, 55) = 5.4, p =0.024).

In summary, the overall aggregate workload score was not different among systems. There were significant differences among the task workload scores, but these differences were not seen among the different systems. The overall aggregate score and average task scores were not

Table 3. Respondents' demographics

	N	Minimum	Maximum	Mean	Standard Deviation
LongUsed (years)	67	0.08	22	6.71	5.82
TimeInGP (years)	67	2	35	17.72	9.06
NumOtherSystms	67	0	6	1.48	1.25
Valid N (listwise)	67				

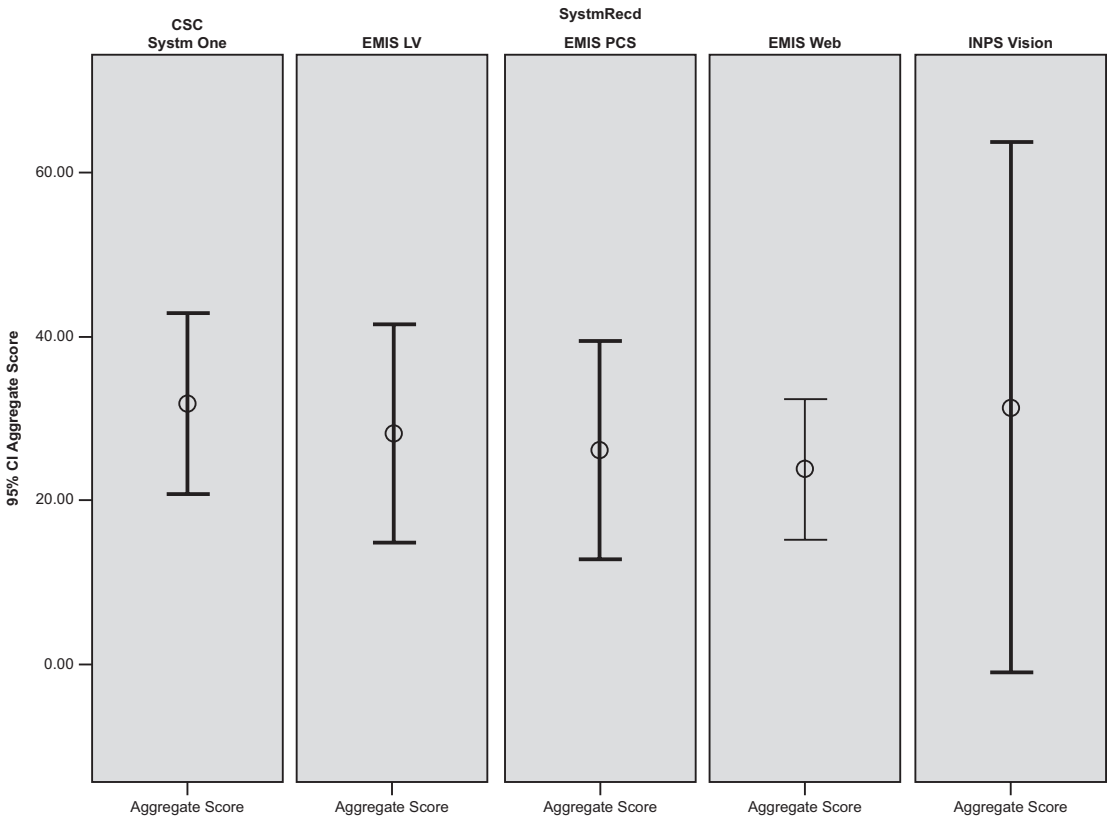

Figure 3. Aggregate score by system

Table 4. Mean scores by task

Task	Mean	95% CI for the mean	Std Deviation
Repeat prescribing	38	31 to 45	28
Find episode	35	28 to 42	28
Overview records	34	27 to 40	26
Drugs management	34	27 to 40	27
Record structured	28	22 to 34	25
Values over time	28	22 to 34	24
Acute prescribing	23	17 to 28	23
Action results	22	16 to 27	23
Record unstructured	21	16 to 26	21
View labs	21	15 to 27	24

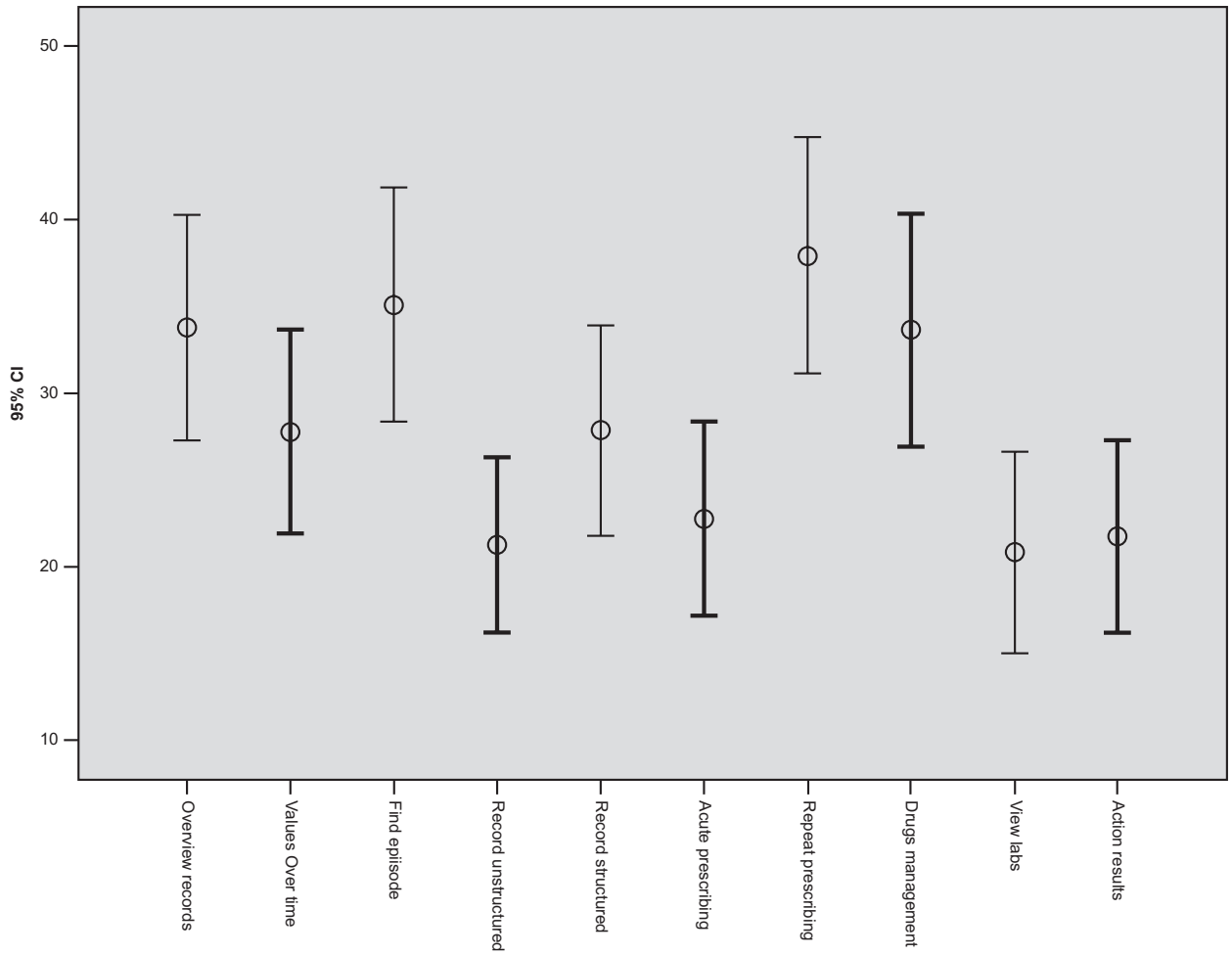

Figure 4. Mean scores by task

related to the time the GPs had been in general practice or the number of other systems used, but they were related to the time the system had been used.

DISCUSSION

Interpretation of results

With the lack of consensus on what should be considered as a threshold for a high or excessive workload[38] – the scores appear to be most useful when used to make comparisons, we looked at the scores in other studies;[22,23,51] compared to these, the workload score found in this study (28.7) appears to sit on the low side, although lower scores have been noticed in some. This probably reflects the familiarity of GPs with their systems and could lead to conclude that no major issues exist.

The tasks evaluated were chosen to represent common tasks routinely carried out by GPs. Tasks with a higher cognitive workload score belonged to the categories 'retrieving information' ('overview records' and 'find episode') and 'prescribing' ('repeat prescribing' and 'medications management'). The task 'overview records' is made of a number of subtasks, but its definition makes sense as a conceptually unique, identifiable task routinely carried out. However, it could be arguably expected that it would incur higher scores than single, smaller tasks. Likewise, the tasks 'repeat prescribing' and 'drugs management' are also composites of smaller subtasks and their higher scores may not come as a surprise either. Prescribing has been an area of high concern in medical errors, and mistakes have been related to the use of computer systems: selecting the wrong drug or dosage instructions from pick lists, overriding drug-drug interaction alerts, unnecessary/inappropriate alerts, the need to maintain an accurate electronic health record and excessive expectations from the computer system.[52] Finally, 'find episode' is a task whose difficulty may indeed be influenced, as many responders pointed out, by the quality and consistency of the recording of the previous episode.

The study did not find any differences on the cognitive workload among the different systems. However, it is likely that the study has been underpowered to detect these. This is discussed further in the following section.

Study limitations

In the study, we introduced a few modifications to NASA-TLX. First, we asked the GPs to recall instead of performing the tasks rather than using alternative methods directly assessing that performance. We considered this logistically necessary to prevent the study interfering with participants' work and negatively impacting responses rates. A study in a nonclinical setting showed that little information is lost when the ratings are obtained retrospectively using a visual recreation of a task,[53] and we found recall worked well in a separate clinical study.[54] Our results should be interpreted in the context of other research that does directly assess performance.[55, 56]

Second, we dropped off the physical demand scale. Modifications involving adding or deleting sub-scales or redefining them have been used in several studies.[38] The specific irrelevance of the physical demand scale when assessing electronic health records has been observed in other studies.[23,54] We proposed that this modification should not have a significant impact on the validity of this study, but undertaking a validity assessment was not within the possibilities of this study.

Finally, we also removed the weighting part. This has been the most common modification made when using NASA-TLX, followed by averaging or adding the scores to compute an estimate of the overall workload (which has been referred to Raw TLX) or analyze sub-scales separately. Studies comparing TLX with Raw TLX have demonstrated the latter to be less, more or equally sensitive; both approaches, however, have demonstrated the diagnostic value of the sub-scales.[38,57]

Like in any survey study where participation is voluntary, this study was subject to a self-selection bias.[58] Also, responses rates in survey studies with physicians are often very low,[59] and electronic means are only of limited utility.[60] The sample we obtained, however, did not appear to differ from the population on a number of variables we had available for comparison. Considering the means and standard deviations observed, sample sizes of around 100 responses per system could have been enough to detect potential existing differences among systems.

Conclusions and implications of the study

Since cognitive workload is only one among many measures of a system's usability, no conclusions on the current GP system's usability quality can be drawn from this study. Also, because of the lack of existence of an acceptable cognitive workload threshold for NASA-TLX in electronic health record evaluations, a sound statement on weather current system's mental workload is acceptable is not possible. Nevertheless, two things can be inferred: first, on the whole, probably no major usability issues exist and second, for some tasks there is probably room for improvement.

The study has identified four GP systems tasks attracting a higher workload score: 'repeat prescribing', 'drugs management', 'overview records' and 'find episode'. Further usability studies could focus further on these tasks, using existing guidance on usability evaluations in primary care.[3,36]

NHS policymakers should remain aware of the importance of usability in system design and might consider following steps in other countries to develop policies around usability requirements.[36]

REFERENCES

1. Goetz Goldberg D, Kuzel AJ, Feng LB, DeShazo JP and Love LE. EHRs in primary care practices: benefits, challenges, and successful strategies. *American Journal of Managed Care* 2012;18(2):e48–54. PMid:22435884.
2. Schade CP, Sullivan FM, de Lusignan S and Madeley J. e-Prescribing, efficiency, quality: lessons from the computerization of UK family practice. *Journal of the American Medical Informatics Association* 2006;13(5):470–5. http://dx.doi.org/10.1197/jamia.M2041. PMid:16799129; PMCid:PMC1561797.
3. HIMSS EHR Usability Task Force. Defining and testing EMR usability: principles and proposed methods of EMR usability evaluation and rating. Healthcare Information and Management Systems Society, 2009.
4. Horsky J, Zhang J and Patel VL. To err is not entirely human: complex technology and user cognition. *Journal of Biomedical Informatics* 2005;38(4):264–6. http://dx.doi.org/10.1016/j.jbi.2005.05.002. PMid:15967732.
5. Peute LWP, De Keizer NF, Van Der Zwan EPA and Jaspers MWM. Reducing clinicians' cognitive workload by system redesign; a pre-post think aloud usability study. *Studies in Health Technology and Informatics* 2011;169:925–9. PMid:21893881.
6. Jha AK, Doolan D, Grandt D, Scott T and Bates DW. The use of health information technology in seven nations. *International Journal of Medical Informatics* 2008;77(12):848–54. http://dx.doi.org/10.1016/j.ijmedinf.2008.06.007. PMid:18657471.
7. Shachak A, Hadas-Dayagi M, Ziv A and Reis S. Primary care physicians' use of an electronic medical record system: a cognitive task analysis. *Journal of General Internal Medicine* 2009;24(3):341–8. http://dx.doi.org/10.1007/s11606-008-0892-6. PMid:19130148; PMCid:PMC2642564.
8. Microsoft Health Common User Interface [Internet], 2010. Available from: http://www.mscui.net/Default.aspx.
9. Patel VL and Kushniruk AW. Interface design for health care environments: the role of cognitive science. *Proceedings of the AMIA Symposium* 1998;29–37. PMid:9929179; PMCid:PMC2232103.
10. Alexander G and Staggers N. A Systematic review on the designs of clinical technology: findings and recommendations for future research. *ANS. Advances in Nursing Science* 2009;32(3):252–79. http://dx.doi.org/10.1097/ANS.0b013e3181b0d737. PMid:19707093; PMCid:PMC3018768.
11. Sittig DF, Kuperman GJ and Fiskio J. Evaluating physician satisfaction regarding user interactions with an electronic medical record system. *Proceedings of the AMIA Symposium* 1999;400–4. PMid:10566389; PMCid:PMC2232602.
12. Nielsen J. Usability 101: Introduction to Usability [Internet]. Nielsen Norman Group, 2012. Available from: http://www.nngroup.com/articles/usability-101-introduction-to-usability/.
13. Levinson W, Lesser CS and Epstein RM. Developing physician communication skills for patient-centered care. *Health Affairs (Millwood)* 2010;29(7):1310–18. http://dx.doi.org/10.1377/hlthaff.2009.0450. PMid:20606179.
14. Pelaccia T, Tardif J, Triby E and Charlin B. An analysis of clinical reasoning through a recent and comprehensive approach: the dual-process theory. *Medical Education Online* 2011 [cited 2013 Jul 11];16. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3060310/.
15. O'Riordan M, Dahinden A, Aktürk Z, Ortiz JMB, Dağdeviren N, Elwyn G et al. Dealing with uncertainty in general practice: an essential skill for the general practitioner. *Quality in Primary Care* 2011;19(3):175–81. PMid:21781433.
16. Starfield B. Primary care. *The Journal of Ambulatory Care Management* 1993;16(4):27–37. http://dx.doi.org/10.1097/00004479-199310000-00006. PMid:10128415.
17. Smith PC, Araya-Guerra R, Bublitz C, Parnes B, Dickinson LM, Van Vorst R et al. Missing clinical information during primary care visits. *Journal of the American Medical Association* 2005;293(5):565–71. http://dx.doi.org/10.1001/jama.293.5.565. PMid:15687311.
18. Beasley JW, Wetterneck TB, Temte J, Lapin JA, Smith P, Rivera-Rodriguez AJ et al. Information chaos in primary care: implications for physician performance and patient safety. *The Journal of the American Board of Family Medicine* 2011;24(6):745–51. http://dx.doi.org/10.3122/jabfm.2011.06.100255. PMid:22086819; PMCid:PMC3286113.
19. Byrne AJ, Oliver M, Bodger O, Barnett WA, Williams D, Jones H et al. Novel method of measuring the mental workload of anaesthetists during clinical practice. *British Journal of Anaesthesia* 2010;105(6):767–71. http://dx.doi.org/10.1093/bja/aeq240. PMid:20846966.
20. Kirsh D. A few thoughts on cognitive overload. *Intellectica* 2000;30:19–51.
21. Clarke MA, Steege LM, Moore JL, Belden JL, Koopman RJ and Kim MS. *Addressing Human Computer Interaction Issues of Electronic Health Record in Clinical Encounters*. Marcus A (Ed). Design, user experience, and usability health, learning, playing, cultural, and cross-cultural user experience [Internet]. Heidelberg: Springer, 2013 [cited 2013 Jul 10], pp. 381–90. Available from: http://link.springer.com/chapter/10.1007/978-3-642-39241-2_42. http://dx.doi.org/10.1007/978-3-642-39241-2_42. PMid:23444324.
22. Ahmed A, Chandra S, Herasevich V, Gajic O and Pickering BW. The effect of two different electronic health record user interfaces on intensive care provider task load, errors of cognition, and performance. *Critical Care Medicine* 2011;39(7):1626–34. http://dx.doi.org/10.1097/CCM.0b013e31821858a0. PMid:21478739.
23. Longo L and Kane B. A novel methodology for evaluating user interfaces in health care. *24th International Symposium on Computer-Based Medical Systems (CBMS)*, 2011, pp. 1–6.
24. Bouamrane MM and Mair FS. A study of general practitioners' perspectives on electronic medical records systems in NHSScotland. *BMC Medical Informatics and Decision Making* 2013;13(1):58. http://dx.doi.org/10.1186/1472-6947-13-58. PMid:23688255; PMCid:PMC3704757.
25. Viitanen J, Hyppönen H, Lääveri T, Vänskä J, Reponen J and Winblad I. National questionnaire study on clinical ICT systems proofs: physicians suffer from poor usability. *International*

Journal of Medical Informatics 2011;80(10):708–25. http://dx.doi.org/10.1016/j.ijmedinf.2011.06.010. PMid:21784701.

26. Viitanen J, Kuusisto A and Nykänen P. Usability of electronic nursing record systems: definition and results from an evaluation study in Finland. *Studies in Health Technology and Informatics* 2011;164:333–8. PMid:21335733.
27. Militello L, Patterson ES, Tripp-Reimer T, Asch SM, Fung CH, Glassman P et al. Clinical reminders: why don't they use them? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 2004;48(15):1651–5. http://dx.doi.org/10.1177/154193120404801505.
28. Garrard CS. Human-computer interactions: can computers improve the way doctors work? *Schweizerische medizinische Wochenschrift* 2000;130(42):1557–63. PMid:11092058.
29. Saitwal H, Feng X, Walji M, Patel V and Zhang J. Assessing performance of an Electronic Health Record (EHR) using cognitive task analysis. *International Journal of Medical Informatics* 2010;79(7):501–6. http://dx.doi.org/10.1016/j.ijmedinf.2010.04.001. PMid:20452274.
30. Zhang J. Human-centered computing in health information systems. Part 1: analysis and design. *Journal of Biomedical Informatics* 2005;38(1):1–3. http://dx.doi.org/10.1016/j.jbi.2004.12.002. PMid:15694880.
31. Khajouei R, de Jongh D and Jaspers MWM. Usability evaluation of a computerized physician order entry for medication ordering. *Studies in Health Technology and Informatics* 2009;150:532–6. PMid:19745368.
32. Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE et al. Role of Computerized Physician Order Entry Systems in Facilitating Medication Errors. *Journal of the American Medical Association* 2005;293(10):1197–203. http://dx.doi.org/10.1001/jama.293.10.1197. PMid:15755942.
33. Middleton B, Bloomrosen M, Dente MA, Hashmat B, Koppel R, Overhage JM et al. Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA. *Journal of the American Medical Association* 2013;20(e1):e2–8. http://dx.doi.org/10.1136/amiajnl-2012-001458. PMid:23355463; PMCid:PMC3715367.
34. Cousins DH and Baker M. The work of the National Patient Safety Agency to improve medication safety. *British Journal of General Practice* 2004;54(502):331–3. PMid:15113513; PMCid:PMC1266164.
35. Protti D. Lessons to be learned from England about the potential of GP computer systems to improve patient safety. *Healthcare Quarterly*2004;7(3):76–80. PMid:15230172.
36. Johnson C, Johnson D and Crowly P. EHR Usability Toolkit: a background report on usability and electronic health records (prepared by Westat under Contract No. HHSA 290-2009-00023I). Rockville, Maryland: Agency for Healthcare Research and Quality, 2011. Report No.: 11-0084-EF.
37. Rubio S, Díaz E, Martín J and Puente JM. Evaluation of subjective mental workload: a comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology* 2004;53(1):61–86. http://dx.doi.org/10.1111/j.1464-0597.2004.00161.x.
38. Hart SG. Nasa-Task Load Index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 2006;50(9):904–8. http://dx.doi.org/10.1177/154193120605000909.
39. Hoonakker P, Carayon P, Gurses AP, Brown R, Khunlertkit A, McGuire K et al. Measuring workload of ICU nurses with a questionnaire survey: the NASA Task Load Index (TLX). *IIE Transactions on Healthcare Systems Engineering* 2011;1(2):131–43. http://dx.doi.org/10.1080/19488300.2011.609524. PMid:22773941 PMCid:PMC3388621.
40. Carswell CM, Lio CH, Grant R, Klein MI, Clarke D, Seales WB et al. Hands-free administration of subjective workload scales: acceptability in a surgical training environment. *Applied Ergonomics* 2010;42(1):138–45. http://dx.doi.org/10.1016/j.apergo.2010.06.003. PMid:20630495.
41. Zheng B, Jiang X, Tien G, Meneghetti A, Panton ONM and Atkins MS. Workload assessment of surgeons: correlation between NASA TLX and blinks. *Surgical Endoscopy* 2012;26(10):2746–50. http://dx.doi.org/10.1007/s00464-012-2268-6. PMid:22527300.
42. France DJ, Levin S, Hemphill R, Chen K, Rickard D, Makowski R et al. Emergency physicians' behaviors and workload in the presence of an electronic whiteboard. *International Journal of Medical Informatics* 2005;74(10):827–37. http://dx.doi.org/10.1016/j.ijmedinf.2005.03.015. PMid:16043391.
43. Hart SG and Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research [Internet], 1988 [cited 2013 Jul 25]. Available from: http://archive.org/details/nasa_techdoc_20000004342.
44. National Aeronautics and Space Administration. NASA TLX: Task Load Index [Internet]. NASA TLX: Task Load Index, no date. Available from: http://humansystems.arc.nasa.gov/groups/TLX/index.html.
45. Laerum H and Faxvaag A. Task-oriented evaluation of electronic medical records systems: development and validation of a questionnaire for physicians. *BMC Medical Informatics and Decision Making* 2004;4:1. http://dx.doi.org/10.1186/1472-6947-4-1, http://dx.doi.org/10.1186/1472-6947-4-18.
46. ISD Scotland. Practice Team Information Annual Update (2009/10) [Internet]. ISD Scotland, 2011. Available from: http://www.isdscotland.scot.nhs.uk/Health-Topics/General-Practice/Publications/2011-02-22/2011-02-22-PTI-report.pdf?.
47. The Health and Social Care Information Centre – Workforce Directorate. General and Personal Medical Services, England, 2001 - 2011, as at 30 September [Internet]. The Health and Social Care Information Centre, 2012 [cited 2012 Dec 15]. Available from: https://catalogue.ic.nhs.uk/publications/workforce/numbers/nhs-staf-2001-2011-gene-prac/nhs-staf-2001-2011-gene-prac-rep.pdf.
48. Davis E. English GP Systems Market [Internet]. Woodcode Consulting, 2011 [cited 2012 Dec 7]. Available from: http://www.woodcote-consulting.com/?p=29.
49. Todd R. CSC to withdraw from primary care. EHI eHealth Insider [Internet]. UK; 2012 Sep 17. Available from: http://ehi.co.uk/news/ehi/8063/csc-to-withdraw-from-primary-care.
50. SystmOne. TPP to become second biggest GP supplier [Internet]. TTP SystmOne Latests News Stories, 2011 [cited 2012

Dec 7]. Available from: http://www.tpp-uk.com/latest-news-stories/tpp-to-become-second-biggest-gp-supplier/.

51. Anders S, Albert R, Miller A, Weinger MB, Doig AK, Behrens M et al. Evaluation of an integrated graphical display to promote acute change detection in ICU patients. *International Journal of Medical Informatics* 2012;81(12):842–51. http://dx.doi.org/10.1016/j.ijmedinf.2012.04.004. PMid:22534099; PMCid:PMC3414670.
52. Avery A, Barber N, Maisoon G, Franklin BD, Armstrong S and Crowe S. Investigating the prevalence and causes of prescribing errors in general practice: the PRACtICe study [Internet]. GMC, 2012 May [cited 2013 Jul 21]. Available from: http://www.gmc-uk.org/about/research/12996.asp.
53. Haworth LA, Bivens CC and Shively RJ. An investigation of single-piloted advanced cockpit and control configurations for Nap-of-the-Earth Helicopter Combat Mission Tasks. *Proceedings of the 1986 Meeting of the American Helicopter Society* (Washington, District of Columbia), 1986, pp. 672–5.
54. Colligan L, Potts HWW, Finn CT and Sinkin RA. Cognitive workload changes for nurses transitioning from a legacy system with paper documentation to a commercial electronic health record. *International Journal of Medical Informatics* 2015;84(7):469–76. http://dx.doi.org/10.1016/j.ijmedinf.2015.03.003. PMid:25868807.
55. Kumarapeli P and de Lusignan S. Using the computer in the clinical consultation; setting the stage, reviewing, recording, and taking actions: multi-channel video study. *Journal of the American Medical Informatics Association* [Internet] 2012 [cited 2013 Jan 17]. Available from: http://jamia.bmj.com/content/early/2012/12/14/amiajnl-2012-001081.
56. De Lusignan S, Pearce C, Kumarapeli P, Stavropoulou C, Kushniruk A, Sheikh A et al. Reporting observational studies of the use of information technology in the clinical consultation. A position statement from the IMIA Primary Health Care Informatics Working Group (IMIA PCI WG). *Yearbook of Medical Informatics* 2011;6(1):39–47. PMid:21938323.
57. Moroney WF, Biers DW and Eggemeier FT. Some measurement and methodological considerations in the application of subjective workload measurement techniques. *The International Journal of Aviation Psychology* 1995;5(1):87–106. http://dx.doi.org/10.1207/s15327108ijap0501_6.
58. Bethlehem J. Selection bias in web surveys. *International Statistical Review* 2010;78(2):161–88. http://dx.doi.org/10.1111/j.1751-5823.2010.00112.x.
59. VanGeest JB, Johnson TP and Welch VL. Methodologies for improving response rates in surveys of physicians: a systematic review. *Evaluation and the Health Professions* 2007;30(4):303–21. http://dx.doi.org/10.1177/0163278707307899. PMid:17986667.
60. Nicholls K, Chapman K, Shaw T, Perkins A, Sullivan MM, Crutchfield S et al. Enhancing response rates in physician surveys: the limited utility of electronic options. *Health Services Research* 2011;46(5):1675–82. http://dx.doi.org/10.1111/j.1475-6773.2011.01261.x. PMid:21492157; PMCid:PMC3207199.