

ANALYZING 10 YEARS OF EARLY AWARENESS AND ALERT ACTIVITY IN THE UNITED KINGDOM

Claire Packer

University of Birmingham

email: c.packer@bham.ac.uk

Matthew Fung, Andrew Stevens

University of Birmingham

Objectives: The aim of this study was to assess the accuracy of the English National Horizon Scanning Centre (NHSC) in identifying and filtering pharmaceutical developments using end user and international collaborator databases of emerging technologies as proxies for new drugs of likely significance to health services and/or patients.

Methods: We used the NHSC information system and the list of National Institute for Health and Clinical Excellence (NICE) technology appraisals to estimate the false positive rate for NHSC identification, filtration, and reporting. We assessed the sensitivity of NHSC identification and filtration of pharmaceuticals for NICE technology appraisals from 1999 to the end of December 2010, and for pharmaceuticals entered into the EuroScan International Network database.

Results: We estimate that overall NHSC identification, filtration and reporting had a positive predictive value of 0.39 (95 percent CI, 0.36 to 0.43) and a false positive rate of 60 percent. Using NICE appraisals and EuroScan's database as proxies for pharmaceuticals of significance, we estimate the NHSC sensitivity over the 10-year period at 0.92 (95 percent CI, 0.89 to 0.95) and 0.89 (95 percent CI, 0.82 to 0.96) respectively.

Conclusions: Our results suggest that the NHSC has performed well in terms of sensitivity over the past decade, but that the false positive rate of 60 percent may indicate that the filtration criteria for pharmaceuticals could be tightened for increased efficiency. Future evaluations of EAA systems should include an element of external review and explore the level of accuracy acceptable to funders and customers of such systems.

Keywords: Horizon scanning, Early awareness and alert systems, Healthcare technology, Program effectiveness

Many countries have established systems to support the uptake of new and emerging health technologies. The National Horizon Scanning Centre (NHSC) in England is one such early awareness and alert (EAA) system, and provides advance notice of new and emerging health technologies and interventions that are likely to have a significant impact on the English National Health Service (NHS) and/or patients, within the next 2 to 3 years (5;8). The work of the NHSC informs the future work program of the National Institute for Health and Clinical Excellence (NICE), the Health Technology Assessment (HTA) program and other national policy- and decision-making organizations.

The NHSC

The NHSC was established as an independent research team at the University of Birmingham in April 1998 and incorporated as a research program within the United Kingdom's National Institute for Health Research (NIHR) in 2006. The Centre has increased significantly in size since 1998 from six to twenty-one whole-time staff with an annual budget of around £1.8M in 2010/11 (US\$2.9M; €2.05M). The Supplementary Table 1, which can be viewed online at www.journals.cambridge.org/thc2012035, sets out the principal features

of the Centre and its methods (available at www.journals.cambridge.org/thc).

Key features of the NHSC methods include, in addition to routine scanning of media and other information sources, collaboration with clinical and technology experts and expert groups, and extensive and proactive contact with commercial companies to identify products in company development pipelines. In addition to its routine scanning and company contacts program, the NHSC undertakes technology- and disease-specific reviews to ensure all clinical areas are scrutinized over time; to focus on areas with known multiple or complex developments, or on areas of particular customer interest; and/or on patient groups with significant or unmet needs.

Public NHSC outputs include short technology briefings and news briefs written using information provided by the company and a defined search of the internet. Other outputs include short confidential filtration forms on emerging pharmaceuticals that are used by customers for initial filtration, and reports from the technology- and disease-specific reviews. The NHSC does not undertake a detailed assessment of any technology or make recommendations for coverage within the health service.

NHSC technology briefings on emerging pharmaceuticals directly enter the decision-making process at NICE and the Advisory Group for National Specialized Services (AGNSS), a group that commissions services for people with very rare disorders. The context for health service decision making around

The authors acknowledge and thank the journal reviewers for their extremely helpful and valued comments. Claire Packer, Matthew Fung, and Andrew Stevens are funded by the United Kingdom National Institute for Health Research (NIHR).

medical devices and diagnostics is much less clear cut, and NHSC news briefs are for wider readership, rather than entering directly into a national decision-making process.

In addition to its main function, the NHSC undertakes research into the development, marketing, diffusion, and impact of new technologies; and the evaluation and development of EAA system processes, methods and outputs. The NHSC is also host to the Secretariat of the International Information Network on New and Emerging Health Technologies (the EuroScan International Network) and a member of the International Network of Agencies for Health Technology Assessment (INAHTA).

Evaluation of Early Awareness and Alert Systems

Although EAA systems have been in place for several years, there has been little published on the efficiency of core methods, for example, identification, filtration, prioritization, and reporting; or the overall system effectiveness or impact on policy decisions or technology diffusion.

Simpson et al. piloted a method normally applied to determine the accuracy of diagnostic tests to estimate the sensitivity and specificity of the NHSC selection and prioritisation process in its first 5–6 years (7). The authors estimated that the NHSC prioritisation had a high negative 0.98 (95 percent confidence interval (CI), 0.92–0.99) and a low positive predictive value 0.14 (95 percent CI, 0.06–0.3); with a sensitivity of 0.71 (95 percent CI, 0.36–0.92) and specificity of 0.73 (95 percent CI, 0.64–0.8).

Douw and Vondeling used a similar method to assess the completeness of identification and accuracy of impact predictions made by oncologists about new anticancer drugs that may impact on Danish health care over the next 5 years (1). Results showed that experts were poor at identifying new anticancer drugs (sensitivity 0.63; 95 percent CI, 0.31–0.86), but were good at predicting which drugs were (positive predictive value 1.0; 95 percent CI, 0.57–1.0) and were not (negative predictive value 0.79; 95 percent CI, 0.52–0.92) likely to have an impact on health care in the future.

Murphy et al. used a Delphi-type process with members of the EuroScan International Network to develop a shared understanding of the important characteristics and components of an effective EAA system (4). The characteristics and components identified as being important for system effectiveness almost all related to structure and process items, with only three related to outcomes.

Packer et al. studied the international diffusion of six new health technologies across 10 developed countries and evaluated the association between diffusion and variables including the presence or absence of early awareness and HTA activity (6). This research found that early awareness and HTA activity, and the presence of a “fourth hurdle” (the requirement to demonstrate clinical- and cost-effectiveness in addition to safety, quality and efficacy) play some part in influencing diffusion of new health technologies. Mundy et al., however, compared the Aus-

tralian pharmaceutical approval and regulatory system (which does not have associated EAA activity) with that of the United Kingdom and Canada (which do have systems to identify pharmaceuticals in development). They concluded that the value of EAA activity is dependent on the systems it is trying to support and that, due to the nature of the system already in place in Australia, an EAA system would neither improve access to new pharmaceuticals nor better inform decision makers (3).

One way of assessing the identification and filtration aspects of a system is to ask if any technologies produced a significant impact on patients and/or health services without first being identified and reported by the EAA system, that is, the system’s sensitivity. However, such an evaluation requires identification of all new technologies deemed to have had a significant impact on patients and health services, before scrutiny of the EAA system outputs. A major problem with this approach is that the value of a technology is not absolute and will be interpreted differently within separate healthcare systems and between health professionals and patients.

Within England, NICE provides guidance to the NHS in England and Wales on the clinical and cost effectiveness of selected new and established technologies. Pharmaceutical topics for technology appraisal come from several sources in addition to the NHSC, including health professionals, the general public, and the Department of Health’s national clinical directors and policy teams. It is therefore possible that if the NHSC did not notify NICE of a key pharmaceutical that was thought to have potential or was impacting significantly on patients or health services, the drug could be notified to NICE from another source and guidance developed. The list of topics appraised by NICE could therefore be used as a proxy for a list of technologies with the potential for significant impact. A second source that could be thought of as a proxy, or at least a comparative source, is the database of new and emerging health technologies of the EuroScan International Network (<http://www.euroscan.org.uk>). Member agencies are encouraged to enter technologies that they have identified. Although identification and filtration criteria and methods differ between agencies, technologies entered are likely to be those with the potential for greater impact than those not identified and entered on the database.

Another approach to assessing the identification and filtration aspects of a system is to estimate the proportion of topics of no significance to patients or health services that were reported to customers—the system’s specificity. Estimation of specificity requires an estimate of the total number of new drugs and new clinical indications for licensed products launched in England over the relevant time period. Using NICE technology appraisals as a proxy for significance to patients and health services, we could then estimate the number of topics launched of little or no significance. However, although it is possible to find out how many licenses to market have been granted from the European Medicines Agency (EMA), some products granted a European license will not be launched in England, and many license

Table 1. Topics Not Reported by the NHSC to NICE, or Not Input Into the EuroScan International Network Database

	NICE proxy	EuroScan proxy
Not in NHSC remit and/or not in filtration criteria at the time of identification	proxy	proxy
Not identified – but not in NHSC remit, e.g., social care, preventive vaccines, food supplements	0	4
Not identified – but unlikely to have fitted filtration criteria at the time e.g. not in remit, not new, ‘me-toos’ in the early years of NHSC activity.	22 + 2*	17
Identified – but not in NHSC filtration/prioritisation criteria at the time e.g. not in remit, not new, ‘me-toos’ in the early years of NHSC activity.	1	7
Topics potentially missed by NHSC		
Not identified – but should have been (potentially within remit and would have met filtration criteria at the time; potential topic we should/would have written at the time).	18 + 3*	8
Identified – but not selected by NHSC filtration. But perhaps should have been selected in retrospect.	0 + 1*	1
Identified – but not thought to be innovative (filtered out). But perhaps was innovative in retrospect.	0	0
TOTAL	44	37

*Three topics required two separate codes for completeness. These were appraisals of multiple drugs, some of which the NHSC had identified, others that had not been; and some identified products that would have fitted filtration criteria and others that would not have done.

variations relate to minor changes to current indications rather than wholly new clinical indications. The number of licenses granted will therefore be an overestimate of the number of new products or clinical indications launch in the England, potentially increasing any estimates of system specificity. Other measures of accuracy include positive predictive value and negative predictive value. Given the comments above we will not be able to estimate negative predictive value in this evaluation.

In this study, we assess the accuracy of the NHSC in identifying and filtering pharmaceutical developments using NICE pharmaceutical technology appraisal topics and EuroScan International Network database entries as proxies for new drugs of key significance.

METHODS

We used NHSC information systems to identify technology briefings sent for consideration of NICE technology appraisal from 1998 to the end of December 2010. We then downloaded a list of all NICE technology appraisals published, currently in development, or proposed until the end of December 2010, from NICE’s Web site (<http://www.nice.org.uk/>). For the estimation of sensitivity around pharmaceutical topics, we excluded technology appraisals of nonpharmaceuticals and reviews of previous appraisals unless a new drug was included in the review, but included topics that were subsequently discontinued. We cross-checked the final list with NHSC outputs to determine whether the NHSC had reported on each new drug before the topic was referred to NICE for appraisal. Appraisal topics without a prior NHSC output were scrutinized independently by CP and MF using an agreed framework to determine why the NHSC had

not informed NICE (Table 1). Any discrepancies between the authors were discussed and agreement reached.

We downloaded all pharmaceutical topics entered onto the EuroScan International Network database between 1 January 2001 and 31 December 2010 from all member agencies except the NHSC. Each entry was cross checked against the NHSC’s internal database to determine whether the NHSC had previously identified the topic. EuroScan International Network database topics that the NHSC had not previously identified were coded independently by CP and MF using the same framework with reconciliation as before.

RESULTS

Since 1998, the NHSC has produced around 950 technology briefings and summaries which entered the prioritisation processes for NICE appraisals (752 topics) or for consideration by the NHSC’s other customers. Of these 295 were products (drugs and non-drugs) that were eventually referred to NICE for appraisal; a positive predictive value of 0.39 (95 percent CI, 0.36 to 0.43) and a false positive rate of 60 percent. Just under 80 percent of the NHSC briefings have related to emerging drugs, the majority of which have been in the cancer field.

NICE Appraisals

We identified 291 NICE appraisals and reviews of appraisals of pharmaceuticals from the NICE Web site to the end of December 2010; with 157 completed appraisals, 116 appraisals in development, and 18 proposed topics. We identified 44 topics without a corresponding NHSC output (table 1). In twenty-three instances, we judged that that the topic would not have fitted the

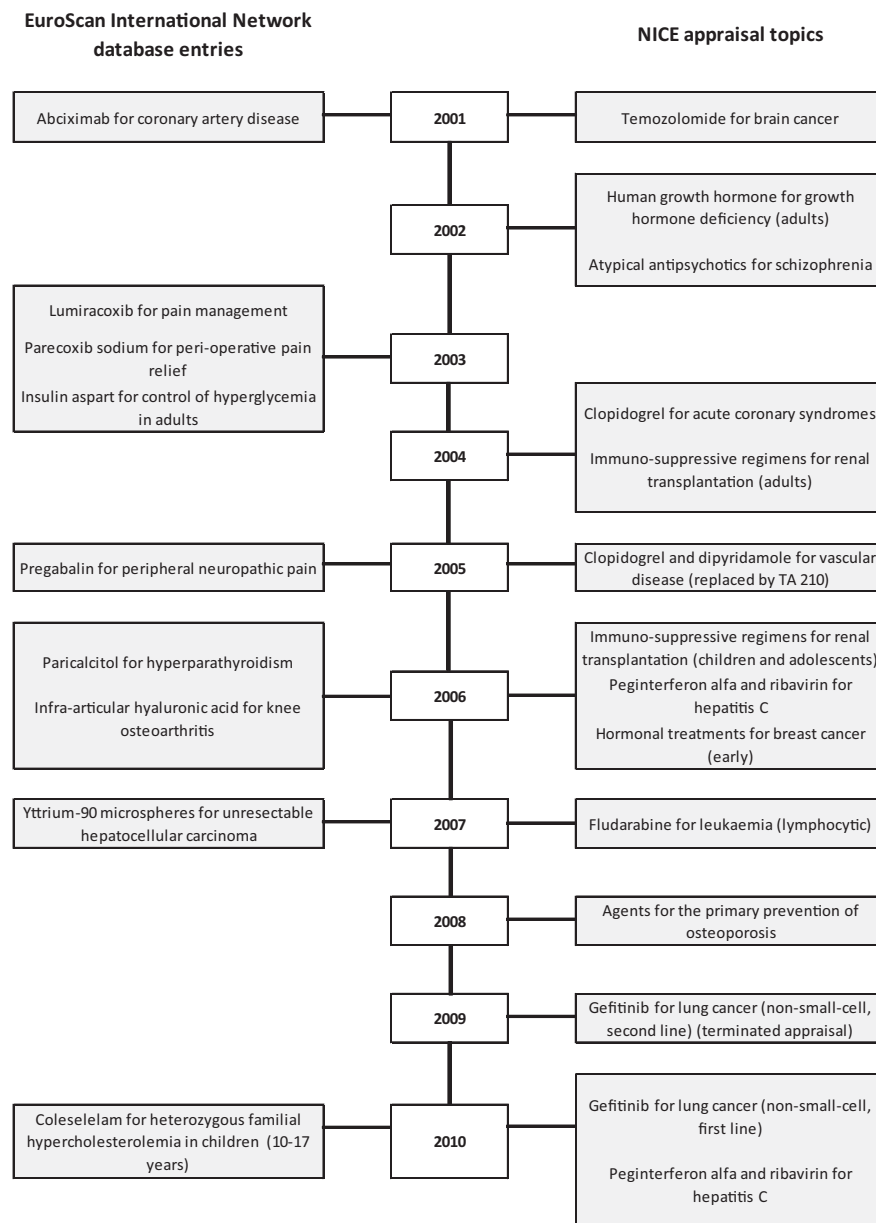


Figure 1. Timeline of topics not reported by the NHSC using NICE appraisal topics and EuroScan International Network entries. Note: Five additional NICE appraisals are either discontinued (and therefore do not have a publication date) or did not have a publication date at the time we undertook the study.

NHSC remit or filtration criteria used at the time, leaving 268 relevant topics. In eighteen instances, we judged that the topics would or should have passed the filtration step at the time, and were, therefore, truly missed. There were also three cases where dual codes were needed, and each of these included a missed topic code. Overall we judged that the NHSC had potentially missed twenty-one (7.8 percent) of relevant topics appraised by NICE, a sensitivity of 0.92 (95 percent CI, 0.89 to 0.95).

EuroScan International Network Database

We identified 108 unique pharmaceutical topics entered onto the EuroScan International Network database from agencies other

than the NHSC to the end of December 2010. Of these, we identified thirty-seven topics without a corresponding NHSC output (Table 1). In twenty-eight instances we judged that the topic would not have fitted the NHSC remit or filtration criteria used at the time. In the remaining nine instances (11.3 percent) we judged that the topics would or should have passed the filtration step at the time, and were, therefore, missed by the NHSC, a sensitivity of 0.89 (95 percent CI, 0.82 to 0.96).

The topics not reported by the NHSC were spread across the decade (Figure 1). Three of the topics were licensed for use very soon after the NHSC was established, and it would have been impractical given the set up time and the relatively small team

to have expected the fledgling NHSC to pick up these “current” topics. Two additional NICE appraisal topics were of multiple drugs of multiple indications in one appraisal topic. The NHSC had identified at least one of the new drugs and/or indications in each case, but not all the products and indications considered in the appraisal.

DISCUSSION AND CONCLUSIONS

Summary of Results

We estimate that NHSC reporting to NICE, for both pharmaceutical and nonpharmaceutical topics, had a positive predictive value of 0.39 (95 percent CI, 0.36 to 0.43). Meaning that almost 40 percent of topics reported to NICE were later selected for NICE technology appraisal and 60 percent were not. Using NICE appraisals and EuroScan’s database as proxies for pharmaceuticals of significance, the NHSC did not report on 7.8 percent and 11.3 percent of technologies that it could or should have identified according to its filtration and prioritization criteria over a 10-year period; sensitivities of 0.92 (95 percent CI, 0.89 to 0.95) and 0.88 (95 percent CI, 0.82 to 0.96), respectively.

Strengths and Weaknesses of the Study

In an ideal world, researchers would be able to evaluate the effectiveness of an EAA system as a whole, rather than evaluating individual aspects of the methods such as identification and filtration. However, there is no such method available and, as the NHSC neither evaluates the technologies it identifies nor issues recommendations to the health service, such methodology may not be relevant in the context within which the NHSC works. A major weakness in the study is that the authors are employed within the NHSC and have been intimately involved in all aspects of the identification, filtration and reporting of emerging technologies to the NHSC customers. This is likely to have introduced some bias, at the very least in interpretation of the results. It would be a positive step for an independent evaluation of an EAA to be undertaken.

A weakness of the sensitivity method used is in the selection of proxy measures of the technologies that have made a significant impact on the English health service and/or patients. Is the list of technologies selected for appraisal by NICE or identified by the EuroScan International Network members a good proxy for technologies having a significant impact on health services or patients? As the NHSC is funded to identify and report on emerging drugs of relevant to the English health service, the list of NICE appraisals is likely to be a better proxy than international activity.

NICE

As any manufacturer, health professional or patient can suggest topics for NICE to consider for appraisal and the development

of guidance, we suggest that any topics of likely or actual significance could be input independently into the system. However, it is possible that this is not a safe hypothesis, and the list of important topics may be much wider than the list of NICE technology appraisals. Indeed, depending on the perspective taken, it is highly likely that individuals will believe that many more new technologies were of importance than NICE selected for appraisal. Although the NHSC may have identified some of these, in this situation we have no easy dataset to compare NHSC activity to.

We also know a small number of topics each year have been identified by the NHSC but not referred to NICE for appraisal, but instead notified to the NICE guidelines development teams for inclusion in NICE guidelines. Some of these may be topics of significance but will not have been captured in the denominator of our comparison.

Given that the NHSC did not identify or select twenty-one emerging drugs for notification to NICE for appraisal, the NHSC’s identification processes may have been insufficient and/or the filtration criteria used may have been too selective. The methods used at both these steps have changed significantly over the years, and the NHSC has widened its scope and broadened its identification sources. It will be interesting to see whether an even greater proportion of future NICE appraisal topics originate from NHSC activities.

EuroScan

As stated earlier the EuroScan International Network database may not be a suitable proxy for pharmaceuticals likely to impact on the English health service, as those member agencies that undertake identification work for pharmaceuticals have differing identification and filtration remits, time frames, and funding. If the agencies’ time frames are shorter (i.e., products are nearer to licensing at reporting) than the NHSC, then you would not expect them to identify emerging drugs before the NHSC, which tries to identify emerging drugs at around 24–30 months before licensing and launch. In this case, this dataset would not be a robust proxy. Additionally, as the NHSC is a major contributor of emerging pharmaceuticals to the EuroScan database, other member agencies who identify emerging pharmaceuticals may not enter the topic onto the database because they assume that the NHSC team will enter it themselves soon (2).

A more accurate proxy of technologies that have made a significant impact would be to elicit opinions from sufficient numbers of health professionals, patients, and health service managers, and commissioners and develop a list of recent technologies that have impacted on the delivery of care, patient experiences and outcomes, and/or health service costs. Depending on the remit given to an EAA system, their output could then be compared with the most relevant opinion-based comparator. Other outcome measures could include looking at the success of NHSC topics notified to research funding organizations for

prioritization for research funds, or for topics that progress through screening program assessments into practice.

Other difficulties encountered in the study include a difficulty in establishing the licensing date of a specific clinical indication in the European Union and the United Kingdom, required to ensure that the NHSC briefing came before licensing, and linking past NHSC briefings to NICE appraisals topics. In some cases the specific patient indication did not correspond exactly between the NHSC topic and later NICE appraisal topic. The objectivity of the authors as assessors of the NHSC systems and activity can also be questioned, and future plans for evaluation of impact and accuracy could include the use of independent assessors.

Relation to Other Literature: Comparative Strengths and Weaknesses

It is tempting to compare sensitivities and positive predictive values found by different evaluations, but this may not be appropriate. Some EAA agencies are more reactive in their identification processes, reacting to individual pharmaceuticals notified to them by clinicians. This is in contrast to the NHSC which is proactive in seeking out all pharmaceuticals in company development pipelines. Being proactive is likely to reduce positive predictive value, whilst being reactive should increase positive predictive value.

Our estimated sensitivity is at the higher end of the range estimated by Simpson et al. and higher than that of Douw and Vondeling. However Simpson et al. were specifically investigating the accuracy of NHSC prioritization processes, and Douw and Vondeling the completeness of expert identification and the accuracy of expert prediction of impact. Our evaluation assesses the accuracy of NHSC identification as well as filtration processes over time. Our estimated positive predictive value of 0.39 (95 percent CI, 0.36 to 0.43), is higher than that found by Simpson et al., but lower than that of Douw and Vondeling. But again the systems being evaluated differ.

Policy Implications of the Findings

Given the legislative context of the NHSC and regulations around the adoption and use of emerging technologies in the English health service, is it ever possible for the NHSC (or other EAA systems) to be 100 percent complete in its identification processes and 100 percent accurate in the application of filtration criteria? Our opinion after working in this field for over 12 years, is no. The only safe method of identifying every possible significant topic would be to require manufacturers to notify the health service or EAA system of every development. However the number of topics likely to make a significant impact will be a sub-set of the total, and identifying those will require resource intensive investigation and filtration, and the application of criteria at some stage in the selection process for more detailed technology assessment and appraisal, and issuance of guidance.

Given the current context, is missing 7.8 percent of emerging pharmaceuticals adequate for the NHSC and its funders and

customers? This is only something that can be decided by the funders and other policy makers and decision makers in the health service.

CONCLUSIONS

We have evaluated the accuracy of the identification and filtration stages of the NHSC and have shown that the NHSC fared well with regards to sensitivity, but not so well with positive predictive value. Reassuringly, the level of missed pharmaceuticals is small when compared with the topics considered important enough to be referred to NICE for appraisal, and to the topics identified and reported by EuroScan International Network member agencies. However, the false positive rate of 60 percent may indicate that the filtration criteria used could be tightened for increased efficiency. Future evaluations of EAA systems should include an element of external review and explore the level of accuracy acceptable to funders and customers of such systems.

SUPPLEMENTARY MATERIAL

Supplementary Table 1

www.journals.cambridge.org/thc2012035

CONFLICT INFORMATION

Claire Packer, BM. BS, FFPH, Director of the NIHR Horizon Scanning Centre, Matthew Fung, BMedSc. MPH, Senior Horizon Analyst, Andrew Stevens, MB. BS, Professor of Public Health, University of Birmingham, School of Health & Population Sciences, Birmingham, United Kingdom

CONFLICTS OF INTEREST

All authors report their institution receives a grant and travel expenses from UK National Institute for Health Research for horizon scanning activity.

REFERENCES

1. Douw K, Vondeling H. Selective new health technologies for evaluation: Can clinical experts predict which new anticancer drugs will impact on Danish health care? *Soc Sci Med.* 2007;64:283-286.
2. Ibarra-Goyen-Roteta N, Gutierrez-Ibarluzea I, Benguria-Arrate G, Galnares-Cordero L, Asua J. Differences in the identification process for new and emerging health technologies: Analysis of the EuroScan database. *Int J Technol Assess Health Care.* 2009;25:367-373.
3. Mundy L, Merlin T, Hiller JE. *Early alert systems for new pharmaceuticals - do they have an impact on pharmaceutical public reimbursement decisions? A cross-national comparison.* Poster Presentation. Globalisation and Health Technology Assessment. 6th Annual Meeting, Health Technology Assessment international (HTAi), Singapore, 21-24 June, 2009.
4. Murphy K, Packer C, Stevens A, Simpson S. Effective early warning systems for new and emerging health technologies: Developing an evaluation framework and an assessment of current systems. *Int J Technol Assess Health Care.* 2007;23:324-330.

5. Packer C. The National Horizon Scanning Centre (NHSC): Early warning for new and emerging health technologies in England. *Evid Based Healthc Public Health*. 2005;9:410-413.
6. Packer C, Simpson S, Stevens A. International diffusion of new health technologies: A ten-country analysis of six health technologies. *Int J Technol Assess Health Care*. 2006;22:419-428.
7. Simpson S, Hyde C, Cook A, Packer C, Stevens A. Assessing the accuracy of forecasting: Applying standard diagnostic assessment tools to a health technology early warning system. *Int J Technol Assess Health Care*. 2003;20:381-384.
8. Stevens A, Packer C, Robert G. Early warning of new health care technologies in the United Kingdom. *Int J Technol Assess Health Care*. 1998;14(4):680-686.