Strategic Intelligence Monitor on Personal Health Systems, Phase 2

Interim Report on Impact Assessment
State of the Art and Justifications

Author: Bernarda Zamora

Editors: Fabienne Abadie, Maria Lluch, Francisco Lupiáñez, Ioannis Maghiros, Elena Villalba, Bernarda Zamora

2013
Table of Contents

1 THE ROLE AND PROMISES OF e-HEALTH..............................................................................3

2 PREDOMINANT MODELS..................................................................................................4
  2.1 Multidisciplinary Models: HTA, MAST and GEMSA ..................................................5
  2.1.1 Health Technology Assessment in the EUenetHTA Model .....................................5
  2.1.2 The Model for Assessment of Telemedicine: MAST .............................................6
  2.1.3 Multidimensional Evaluation Grid for Health and Autonomy: GEMSA ..................9
  2.2 Randomised trials: Whole Systems Demonstrator ....................................................10
  2.3 Systematic reviews and other models .....................................................................11
  2.4 Reasons for the lack of conclusive evidence .............................................................13

3 NEEDS FOR EVIDENCE ON COST-EFFECTIVENESS AT MACRO LEVEL ...........13
  3.1 Limitations of the current approaches ...................................................................13
  3.2 Micro-macro models on the impact of telehealth and telecare ................................15

4 CONCLUSIONS ..............................................................................................................16

List of Figures
  Figure 1: Hierarchy of evidence ....................................................................................4
  Figure 2: Framework of Evaluation in the health ecosystem .........................................5

List of Tables
  Table 1: Headline benefits of eHealth, as discussed by SDO ........................................4
  Table 2: Reconciliation between GEMSA and MAST ..................................................9
1. THE ROLE AND PROMISES OF E-HEALTH

The challenges for healthcare systems in the European Union include demand side and supply side drivers. On the demand side, demographic changes due to ageing and increasing personal income are shaping growing expectations on healthcare services to increase quality and access, reduce disease burden, respond to emergency disease risks, and assist mobility and adaptation to the workplace. On the supply side, healthcare systems are under the pressure from limited budgets and the increasing complexity of healthcare provision which requires the management of investments in technology and interoperability of information flows alongside organizational changes.

In the context of the Europe 2020 strategy and in order to answer these challenges, the European Commission has launched the pilot European Innovation Partnership (EIP) on active and healthy ageing (AHA) whose overarching target is to increase by two years the average healthy lifespan of European citizens by 2020. eHealth systems are expected to play a key role in achieving all of the three overall objectives of the EIP-AHA initiative which are to have a positive impact on a) the health status and quality of life of the EU ageing population, b) on the sustainability and efficiency of health and social care systems, and c) on the competitiveness of the EU industry, especially in the Health and ICT sectors.

The increase in life expectancy can occur under different scenarios depending on the health status of individuals during the Last Years of Life. The potential contribution of eHealth and IPHS on the management of chronic diseases is discussed in Codagnone et al. (2011). However, since chronic diseases cannot be reversed, one should expect that better management by using IPHS would contribute to extend life but may have perverse consequences in terms of higher prevalence and health costs. This negative view does not take account of all the potential benefits affecting directly hospitalisation costs and a better use of resources and health and social care professionals, and indirectly through positive cross-sectoral spillovers: by increasing work productivity of relatively younger cohorts (the 55-64 age group), and by increasing the contribution to GDP growth of innovative industries in the sector and also their contribution to tax revenues.

As shown in Table 1 the NHS Service Delivery and Organisation (SDO) programme has identified the possible contributions of eHealth. In particular, those items which fall under the heading of IPHS, and of the broader defined eHealth applications have been extensively discussed in the literature. The most remarkable contributions of eHealth are derived from its impact on information sharing, improvement of clinical accuracy for diagnosis and treatment, support of prevention activities, and staff and workload organisation. The overall effect of improving health status as a component of the human capital of a society is in itself a factor both of technological growth and labour supply in a country’s macroeconomic production function.

---

Table 1: Headline benefits of eHealth, as discussed by SDO

<table>
<thead>
<tr>
<th>Improving healthcare <strong>access</strong> by:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• helping to alleviate barriers to effective healthcare introduced by physical location or disability</td>
<td></td>
</tr>
<tr>
<td>• facilitating consumer empowerment for self-care and health decision-making.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Improving healthcare <strong>quality</strong> by:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• aiding evidence-based practice</td>
<td></td>
</tr>
<tr>
<td>• tailoring care to individuals, where IT enables more informed decision-making based on evidence and patient-specific data</td>
<td></td>
</tr>
<tr>
<td>• improving transparency and accountability of care processes and facilitating integrated and shared care</td>
<td></td>
</tr>
<tr>
<td>• reducing errors and increasing safety.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Improving healthcare <strong>cost-efficiency</strong> by</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• streamlining healthcare processes, reducing waiting times and waste</td>
<td></td>
</tr>
<tr>
<td>• improving diagnostic accuracy and treatment appropriateness.</td>
<td></td>
</tr>
</tbody>
</table>

*Source: Pagliari C. et al. (2004)*

## 2. PREDOMINANT MODELS

Numerous studies have been undertaken to evaluate the impact of telecare and telehealth. Figure 1 presents a ranking of methodologies which have been used in these studies. The most recent evaluation models, still in design and testing phases, combine several of these methodologies, including experts’ opinion, but with the advantage of doing so under a holistic evaluation framework which adapts the method to the dimension to be evaluated. The most recent evaluation models are the multidisciplinary models MAST and GEMSA which are analysed below. Given their multidisciplinary approach, several methodologies are used, including systematic reviews. Also, a recently concluded large scale pilot run in England, the Whole Systems Demonstrator (WSD), follows an evaluation model of randomised trial but with a more pragmatic and comprehensive focus than traditional Randomised Controlled Trials.

**Figure 1: Hierarchy of evidence**

*Source: Davies and Newman WSD Action Network briefing paper 2011.*

---

2.1 Multidisciplinary Models: HTA, MAST and GEMSA

A multidisciplinary model is expected to evaluate the impact of eHealth within the framework of the health ecosystem as shown in Figure 2. The following summaries of the HTA, MAST and GEMSA models are compared with the health ecosystem to identify the contributions and limitations of these methodologies to the broadest field of health economic evaluation. At the end, one should expect that the new models bridge the gap between medico-clinical evaluations and economic evaluations, including micro-to-macro approaches.

Figure 2: Framework of Evaluation in the health ecosystem


2.1.1 Health Technology Assessment in the EUnetHTA Model

The Health Technology Assessment Core Model (HTA) covers the broad field of multidisciplinary analysis of medical, social, economic and ethical aspects of the impact of technology in health. The European Network for Health Technology Assessment (EUnetHTA) has developed the HTA Core Model (2008) applied to Diagnostics technologies on the one hand and to Medical and Surgical interventions on the other. The methodology includes ten domains which can be included in the framework presented in Figure 2.

The health problem and current use of technology and the description and technical characteristics domains are preceding considerations needed to define blocks A to D of the framework. These domains describe the intervention to be assessed in terms of epidemiological problem, target population and procedures, equipment and resources used in the status quo versus the new technology to be applied. The three domains related to clinical outcomes, which could be included in block A of the above figure, are Accuracy, Clinical Effectiveness and Safety. The evaluation of

---

clinical outcomes feeds from the methodology used to test new drugs in the pharmaceutical industry, mainly Randomised Controlled Trial (RCT).

The *economic evaluation* is captured in a new domain intended to inform decision-makers on different value-for-money judgments. If the economic evaluation is limited to assess a particular intervention at provider level, the methods are those of the micro evaluation (costs-effectiveness, cost-benefit, cost-utility). On the other hand, if the intervention covers providers of different regions or countries, the need for comparison and transferability across different organisations and healthcare systems, requires a macro evaluation as well as the consideration of spillover effects across sectors. In the economic domain, the most common macro indicators refer to equity, efficiency, labour productivity, economic growth, and fiscal impact.

An important domain considered in the HTA Core Model but which is not emphasized in the health ecosystem in Figure 2 is the focus on service delivery models captured in the *Organisational* domain. It has been acknowledged that regulations and procurement rules in particular, as well as differences in reimbursement models have an important effect on the deployment of technologies in healthcare. Therefore, this domain constitutes a separate domain which is given a particular focus in all of the ongoing multi-dimensional evaluation models.

Lastly, *Ethical* and *Social* aspects constitute two separate domains related to culture. Even though the organisational, ethical and social domains are broader than the *Legal* domain, they are interlinked since national and international legislation establish standards on safety, patients’ rights, competition, and other issues determining the important aspect of liabilities.

One can conclude that the health ecosystem framework presented in Figure 2 is incomplete in the sense that it only includes all these aspects either as "non-healthcare determinants" or as "spillovers" while the last four HTA domains mentioned above characterise the proper institutional context of the health system.

### 2.1.2 The Model for Assessment of Telemedicine: MAST

The Model for Assessment of Telemedicine (MAST) is the outcome of the EC funded MethoTelemed Project carried out between February 2009 and February 2010. The project was led by MedCom, Denmark, in partnership with the Norwegian Centre for Integrated Care and Telemedicine. The MethoTelemed project started with a multidisciplinary focus by systematically reviewing and analysing research on telemedicine to identify the different assessment and validation methodologies developed to date. Moreover, the project integrated stakeholders views obtained during two workshops held in June and November 2009.

The MAST and EUnetHTA models are equivalent in terms of multidisciplinary focus and aspects covered: medical, social, economic, and ethical. However MAST aggregates and sequences these aspects differently by considering seven domains of outcomes measuring effectiveness and contribution to quality of care of telemedicine applications. MAST uses the core model and terminology of the EUnetHTA model and aims to establish a common model for assessment of telemedicine in EU Member States.

The common domains between MAST and EUnetHTA are *Safety, Clinical Effectiveness, Economic, Organisational, Socio-cultural, ethical and legal aspects*. On the domain definition, the novelty of MAST lies in the *Preceding Considerations* block and a last block to assess *Transferability*. Also, within the two EUnetHTA descriptive domains of the *health problem and current use of technology* and the *description and technical characteristics of the technology*, MAST makes a distinction between those aspects which are relevant for the preceding considerations (purpose, alternatives, geographical scope and size, maturity) and those which are the object of the evaluation (*health problem and characteristics of the application*). Lastly, MAST defines a new specific domain called *Patient Perspectives* which captures patients’ perceptions which can be considered either under the

---

10 EUneHTA, 2008, HTA Core Model for Medical and Surgical Interventions: [www.eunethta.net](http://www.eunethta.net)
social or the ethical domains in the HTA Core Model (EUnetHTA). The rationale of MAST is to develop a new model to fill the existing gap for a coherent and reliable set of evidence of the effects of telemedicine. This new model was presented in February 2010 as the MAST Manual and MAST toolkit.11

To the best of our knowledge, the MAST methodology is being applied and tested in the RENEWING HeALTH project (Regions of Europe Working Together for Health), which is partially funded by the European Commission under the ICT Policy Support Programme as part of the Competitiveness and Innovation Framework Programme and running from February 2010 to September 2012. This project is being implemented jointly in nine regional authorities of nine EU Member States (Italy, Denmark, Sweden, Norway, Spain, Finland, Greece, Austria, and Germany). The project covers 18 studies of telemedicine applications, including a total of 5,000 patients suffering from one or several of the three main chronic diseases (diabetes, COPD, CHF). All of the 18 pilots follow a common design based on a pragmatic RCT with a follow-up period of 18 months. Moreover, a real life pilot is planned with the aim to compare the assessment under the ideal conditions of the RCT (measuring efficacy) and real life conditions where outcomes are affected by different confounders (measuring efficiency). After the completion of the assessment on transferability, the final aim of the project is to provide decision makers with both a self-evaluation tool (a standard Telemedicine Readiness Index) and expert advice for the implementation of large scale telemedicine services. This must be reported in a Deployment Plan depicting a roadmap for full deployment in the whole territory of the nine participating regions. At national level MAST is being used in Denmark in the ‘Patient Briefcase’ (COPD trial) initiative and in the ‘Diabetes foot ulcer’ project (See SIMPHS2 Country Study Denmark).12 For the ‘Patient Briefcase’, the patient inclusion process is closed and evaluation will start in April 2012. In addition, as reported by experts at the SIMPHS2 Validation workshop12, France also asked to adapt MAST for French telemedicine projects and there have been contact with SMEs wanting to use MAST for mHealth application. It is hoped that positive outcomes will be reached by using this method.

On February 10, 2012, the pilot teams convened in a MAST seminar13 with experts in the different MAST domains to agree on a guideline for constructing comparable tables and data analysis across countries. This guideline defined (i) the protocols; (ii) the minimum dataset, including demographics, health related quality of life (from the SF-36 health survey which is composed of 36 questions), patient acceptability (from the Whole System Demonstrator questionnaire), economic aspects, and organizational aspects; and (iii) the minimum level of analysis.

Before carrying out an assessment, the Preceding Considerations block must determine whether the application is mature enough to allow for the definition of outcomes and the assessment of effectiveness and quality through a summative evaluation. If the application is not mature, only safety and feasibility have to be analysed. Also, the fixed costs and economies of scale must be considered in defining the level of deployment or sample size in a pilot project.

The evaluation of the seven domains is composed of qualitative and quantitative indicators. On the qualitative side, the assessment includes, as preceding considerations, a general description of the health problem and the clinical practice to treat it, the purpose, technical characteristics, and requirements of the telemedicine intervention. Equally, given the need to understand the institutional context, the organizational, ethical, social and legal domains require an evaluation of qualitative nature. The domain of safety (clinical and technical) must also be described on the basis of industry and R&D information, as well as whether there is any type of accreditation for commercial release.

On the quantitative side, several quantitative indicators reflect the health condition, the patient’s perception, and the economic aspects to inform value-for-money judgments. Besides a qualitative

11 The MAST Manual and MAST toolkit can be downloaded at www.telemed.no/methotelemed
12 http://is.jrc.ec.europa.eu/pages/TFS/SIMPHS2.html
description, the health problem also has a quantitative assessment in terms of measuring disease burden collected from demographic data and health-related quality of life data (e.g. prevalence, incidence, mortality, disability). The domain of clinical effectiveness (effects on the patient’s health) collects quantitative statistical measures of effects under ideal circumstances (efficacy), and in regular clinical practice (effectiveness). This is achieved by designing pilots in both randomized controlled and pragmatic trials, where the patients under the intervention are compared with a control group and the effect of demographic characteristics, including co-morbidity scores, is taken into account.

The patient perception new domain is quantified using the Service User Technology Acceptability Questionnaire (SUTAQ) designed by Professor Newman for the Whole System Demonstrator. The MAST patient perception dimensions do not match exactly the SUTAQ dimensions, but several analyses, including multivariate regression on the SUTAQ scores can be performed to measure global satisfaction and acceptance. Access and ability to use the application can be directly measured according to SUTAQ dimensions. And confidence, empowerment and self-efficacy can be approximated by questions related to privacy and care personal concerns.

The economic evaluation is also a quantitative assessment which recommends measuring both costs-effectiveness (costs related to a single common effect) and cost-utility (measuring cost per healthy year or cost per quality-adjusted life-year) at the level of the healthcare provider, the municipality and the patient. This economic evaluation follows the guidelines of Drummond et al. (2005). Data on costs (activity, use of time and prices) can be collected by interviewing staff from the health care provider and through surveys of providers and patients. The SF-36 health survey can be used to measure outcomes in terms of health status (QALYs) The results are presented in terms of average costs of treatment per patient in the treated and control groups and in terms of incremental cost-effectiveness ratio (ICER) which measures the cost of one additional unit of effect by comparison of the treated and control average costs and effects. This economic evaluation is therefore a Micro-evaluation and does not cover macroeconomic aspects and cross-sectoral effects such as labour productivity or fiscal impact.

To assess organizational aspects, the main part consists in a qualitative evaluation of the work flow and patient partway, and the structure and culture within the provider organization. Complementarily, some data are presented to describe some of these elements such as number of staff and distribution, visits to GP amongst others. Also on the qualitative side, the ethical domain reports on wider ethical issues, such as patient autonomy, access and equity and the existence of an ethical review. These aspects also appear in the legal domain to inform on liabilities and responsibilities, and confidentiality (data protection). The social and cultural aspects deal with changes in roles and responsibilities for patients and their relatives, including gender aspects. The key outcomes contain qualitative information which can be quantified if collected through patient’s surveys.

The assessment of transferability answers under what conditions the particular results obtained after the evaluation of a telemedicine application can be replicated and how much these results are expected to change when the application is applied (i) in other countries (cross-border transferability), (ii) at a larger scale (scalability), and (iii) for different types of patients (generalisability). The key aspects of the setting which determine the trial results and their external validity should be described in the safety and clinical effectiveness domain, and these must be compared with alternative settings in the organisational domain jointly with other types of variables for which a confounder effect has been found in the trial (e.g. demographic composition).

The main limitation of the MAST model is the lack of an analysis of processes and of the causation mechanisms which hinder the analysis of transferability.

### 2.1.3 Multidimensional Evaluation Grid for Health and Autonomy: GEMSA

The GEMSA multidimensional model is presented in Le Goff-Pronost and Picard (2011) as a new multidimensional evaluation model for evaluating telehealth. The authors argue on the need to add new dimensions to the traditional evaluations based on clinical and medico-economic evaluations. Equally to the MAST model, the foundations of GEMSA are in HTA models. Despite the broad approach of HTA and the existence of HTA guidelines for several years, there is no agreed matrix of outcome indicators to evaluate telehealth yet, neither is there an accepted definition for those that exist.

The first contribution of the GEMSA model is on the definition of outcome indicators which complete the set presented in the Canadian project “National telehealth outcome indicators” (NTOIP). The aspects included in NTOIP were: quality, access, acceptability and costs. The aspects of quality, access and acceptability can be matched with indicators of the MAST method described before, in particular in the domains of safety, clinical effectiveness and patients perspectives. Cost is an aspect pertaining to the economic domain.

Even though MAST includes a broader range of criteria, it has not offered a clear set of outcome indicators and reference values yet. In this sense, the GEMSA model contributes to filling the gap between multi-criteria and standardisation, first because its five general categories cover more criteria than the four included in NTOIP so as to complete all the aspects of a full HTA model, and second because it matches the five general categories with outcome indicators which are measurable either by objective or by self-evaluation tools.

GEMSA presents the five categories in two radar graphs, one for an instruction grid to assess feasibility of a proposed project, and the other one for an evaluation grid to assess the stages of the project during execution. The five categories are: (i) Strategy which covers the purpose, method and contribution of the project. These are aspects defined under three domains in MAST; (ii) Technology, technological and industrial expertise which cover innovation, standards for production and use, and safety; (iii) Organization which describes benefits for users in the fulfilment of their missions; (iv) Quality which defines benefits for users in terms of their needs; and (v) Economics which assess economic viability and whether the project generates new economic activity. Table 2 shows the domains that are common to both the GEMSA and the MAST methods even though the indicators collected by each model may be different.

#### Table 2: Reconciliation between GEMSA and MAST

<table>
<thead>
<tr>
<th>MAST Domain</th>
<th>GEMSA category</th>
<th>Health problem and characteristics of the application</th>
<th>Safety</th>
<th>Clinical Effectiveness</th>
<th>Patient Perspectives</th>
<th>Economic</th>
<th>Organisational</th>
<th>Socio-cultural, ethical and legal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategy</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality and usage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>(GEMSA broader than patient)</td>
<td></td>
</tr>
<tr>
<td>Economic aspects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Source: Le Goff-Pronost & Picard (2011)

---


GEMSA is currently used by a reference centre in France to complete an evaluation file for common strategic initiatives. In this test of the GEMSA model, a scientific committee evaluates five outcomes for each one of the five categories, and included a sixth category of project management corresponding to the instruction grid. The outcomes are scored from 1 to 5, adding score zero for elimination. Categories *Organization* and *Economics* weight half the rest. The most determinant test to validate the GEMSA model has been carried out as part of a European project, Interreg IVP “Suode” where the evaluation of telemedicine is a specific task. The robustness of the GEMSA categories has been demonstrated since different experts from the evaluation committee have converged on a set of fifteen sub-categories, three in each one of the five GEMSA categories. The consensus achieved guarantees that GEMSA can be used as a common practice to evaluate and compare different types of telemedicine projects.

### 2.2 Randomised trials: Whole Systems Demonstrator

The Whole Systems Demonstrator (WSD) is a large scale pilot for telemonitoring patients with long-term conditions (Diabetes, COPD and CHF) and social care needs funded by the English Department of Health. Around 6,000 patients have been selected in Primary Care Centers at three sites (Cornwall, Kent and Newham). The evaluation method used "[...]

Again, the WSD trial aims at providing rigorous evidence on impact of telehealth and telecare on quality of care and reduction of costs and at the same time have the support of different stakeholders. To have this support, the randomisation method used has been based on clusters defined at GP practice levels instead of traditional RCT at individual patient level. Therefore, the WSD also has a multidisciplinary focus structured around five domains or themes: (i) service utilization; (ii) clinical effectiveness; (iii) cost-effectiveness; (iv) patient and professional experience; and (v) service delivery and organization. The first three themes are the focus of the core quantitative analysis based on the RCT methodology, and the two last themes are analysed separately by qualitative methods.

The number of outcomes in the WSD is smaller than in the multidisciplinary models presented above. For service utilization, a unique outcome on use of hospital services is defined based upon existing data from hospital databases. For clinical and cost-effectiveness, a questionnaire is filled by patients and carers to report on different outcomes of patients’ and carers’ perception, and health utilisation data are collected via interviews. The patients in the intervention group use either telehealth or telecare, while the patients in the control group receive usual care. There are separate analysis for telehealth and telecare.

Cost estimates for the intervention and control group are derived from multivariate and univariate analyses of costs computed at four different levels (National Health Service (NHS), Local authority, NHS and Local authority, and the broader public sector). Benefits for patients are measured in QALY and through the ICECAP capability index. The economic outcomes indicators are ICER and the cost-effectiveness acceptability curve (assessing in terms of probability). In this case, these indicators are cost-utility measures since cost are compared with quality adjusted life indicators. On clinical effectiveness, besides quality of life, some other indicators on well-being, self-care, and carer burden are presented. All outcomes are measured at three time points: baseline, 3 months and 12 months.

---


2.3 Systematic reviews and other models

Systematic reviews have been considered as the highest quality evidence in the sense of informing on the most accurate effect. In the case of telehealth and telecare, most of the existing systematic reviews are based on a review, selection and summary of scientific literature. There exist some systematic reviews performing meta-analysis based on the Cochrane Handbook of Systematic Reviews for Interventions.\textsuperscript{20} However, given the lack of consistent and robust evidence found in most of the cases, there has not been an effort to explain the variability of results by using meta-regression. Such meta-regression has been applied for instance to analyse the impact of hospital ownership on health outcomes in Eggleston et al. (2008)\textsuperscript{21} using random-effects meta-regression analysis to quantify to what extent various study characteristics account for heterogeneity of findings in the literature.

A recent systematic overview and synthesis of the literature on the impact of eHealth on the Quality and Safety of Healthcare is presented in the Final Report for the NHS Connecting for Health Evaluation Programme\textsuperscript{22} (NHS CFHEP). The report aims at supporting the National Programme for Information Technology (NPfIT) which was launched by the NHS in 2002 and planned to implement the longitudinal patient record (NHS Care Records Service) together with national network and databases, an electronic prescription system, and an electronic appointment booking service (Choose & Book). More recently, the Programme expanded to include telehealth and telecare initiatives. The report covers all areas in eHealth divided into three domains relating to key activities they support, included IPHS which is described by the term “telehealthcare” and refers to “the provision of personalised healthcare over a distance” (by means of IT).

The outcomes chosen to measure impact of telehealth and telecare in the NHS CFHEP are divided into quality domains, and safety whose taxonomy follows the Joint Commission of Accreditation of Healthcare Organizations (JCAHO). The report includes systematic reviews and meta-analysis from the Cochrane Library, from searches through MEDLINE and EMBASE, and from the author’s personal databases. A total of 162 systematic reviews are referenced and a summary sheet of each one is included. Some of these studies cover eHealth at the level of Electronic Health Records, ePrescription and Diagnostic, and the report also includes around 30 systematic reviews and meta-analysis about telehealthcare as “remote delivery of care”. Several studies cover randomised interventions for several diseases\textsuperscript{23} 24 25 26 27 28 29 30 with a variety of outcomes and sometimes

\begin{itemize}
\item Higgings, J. and S. Green (2011) \url{http://www.cochrane-handbook.org/}.
\end{itemize}
inconclusive evidence, although a meta-analysis on several interventions for health disease, diabetes and psychiatric conditions indicates that telehealth positively affects clinical outcomes of care, even in different patient populations. The report also includes systematic reviews focused on one condition. For CHF, two meta-analyses conclude that telehealthcare (telephone support or home telemonitoring) is effective in reducing the risk of all-cause mortality and CHF-specific hospitalisation. However, a third meta-analysis concludes a non-significant effect on reduced all-cause mortality. In the case of diabetes, the meta-analyses included find either a non-significant effect of telecare on glycemic levels or a favourable effect, pointing also towards reduced hospitalization. Lastly, in the case of respiratory diseases (asthma) benefits of telehealthcare seem limited to reduced risks of hospitalisation for patients in high risk group. On the basis of the included systematic reviews of cost effectiveness studies of telemedicine interventions, there is no relevant evidence to confirm whether or not telemedicine is a cost effective alternative to standard healthcare delivery.

Another methodology consists in the analysis of several case studies. This has been used in some evaluation projects covering different countries of the European Union and the broader field of eHealth, including Electronic Health Records. This is the case of the EHR Impact and eHealth Impact projects. For the case of telehealth and telecare in particular, a large study was conducted by Gartner for the Ministry of Health and Social Affairs in Sweden during the Swedish Presidency of the European Union in 2009. The study was based on data obtained at 300 clinical data points in six member states, and included telemedicine among other eHealth technologies. This study found indicators on 37 types of benefits classified in five domains. The methodology to assess impact is based on experts self-reporting and on extrapolations of impacts found in the literature (e.g. on hospital admissions).

Prior to evaluation projects, the European Commission embarked on a broad eHealth benchmarking project which may be considered and even used to carry out ex-ante evaluations of telehealth and telecare since telemedicine/monitoring was included as one of the activity-dependent indicators. The broad benchmarking surveys collected data on use of telemonitoring and social alarms by

40 Gartner (2009) eHealth for a Healthier Europe –opportunities for a better use of healthcare resources.
informal carers, GPs, and nurses. Although the indicators related to impact, on general self-perceived attitudes, captured attitudes and analyzed best practices at global eHealth level, this benchmarking project mapped the status of eHealth and its use in EU in 2002 and 2007, when the surveys were collected. There is still a potential for using these activity data for telehealth and social alarms in new evaluation projects, especially to measure activity at the status quo point versus the intervention, and to assess scalability.

Lastly, it is worth to mention the broad proposal of Social Impact Analysis of ICT on Health\textsuperscript{42}, funded by the European Commission, which presents a different framework along four domains: (i) Rationalization, (ii) Networking and Social Capital, (iii) Empowerment and Participation, and (iv) Information and Lifelong Learning. The domain of Rationalization summarises some evidence on impact of telehealth, even on ethical issues. This domain is common to the current impact evaluation methodologies (MAST, GEMSA and WSD). The remaining three domains are too blurred and focus mainly on patients. Moreover, the report mainly discusses these aspects applied to the use of internet to search for health information ("health-on-the-web").

2.4 Reasons for the lack of conclusive evidence

The systematic reviews on studies about telehealth impact are numerous, some of them very recent and which overview systematic reviews carried out during the last decade. The most recent systematic overview presented in the above mentioned final report NHF CFHEP analyzes studies published during the period 1997 to 2010. Most of the studies referred to the impact of telehealth covering small scale projects with less than 500 patients. The methodology used to measure impact is sound, RCT in most of the cases, but it is limited to measuring only a few clinical indicators on quality such as risks of hospitalisation or mortality. Costs are not well measured since the time frame considered is short and there is no evaluation of opportunity costs. Therefore, the mere scale and particularity of the cases analyzed is the main reason for the lack of conclusive evidence and even more for the consideration of this evidence at a larger scale.

The design of multidisciplinary models such as MAST and GEMSA, and also the WSD as large scale project, has followed a rationale to achieve conclusive evidence. Therefore, these models have been designed taking into account the reasons for the lack of conclusive evidence on the impact of telehealth and have focused on improving three main factors: (i) definition of outcomes and measures in a multidisciplinary setting; (ii) robust theoretical framework to design the interventions with due consideration to evaluation; and (iii) definition of the proper time frame to measure the outcomes.

3. NEEDS FOR EVIDENCE ON COST-EFFECTIVENESS AT MACRO LEVEL

3.1 Limitations of the current approaches

The new evaluation models for telehealth, especially multidisciplinary models, have emphasized the need to evaluate costs relative to effects and/or benefits and have considered a domain to include indicators for economic evaluation, or costs-effectiveness in the case of the WSD. However, even though MAST proposes an Economic Evaluation with a societal perspective, the principal focus on the measurement of costs and use of resources is a provider perspective based on microeconomic indicators collected at the healthcare provider site (e.g. investments in equipment, staff, outpatient and primary care visits, hospitalisation). On the side of effects, the model limits the proposal to clinical effects. This economic evaluation model is preserving the type of partial and case-based

economic evaluations which have been carried out to study specific small scale telehealth projects (see recent systematic overviews\textsuperscript{43,44}).

At the level of the healthcare system or institutional level, the economic evaluation method proposed in MAST is to assess the business case in terms of Returns on Investment (ROI), with the added difficulty of how to assess the revenue or reimbursement for telehealth given the lack of specific Diagnostic Related Group (DRG) to classify and determine the average costs (reimbursement rate) for telehealth treatments. The GEMSA methodology contains a general question to define the “Economics” dimension about whether the project is economically viable and whether it generates new economic activity.

Yet, economic viability can confront the views of the industry and health care providers with those of the overall health care system. The industry and healthcare providers can assess the economic viability of their business cases in terms of ROI and risk, but it is the responsibility of the governments to assess the fiscal viability and the burden on public budgets and societal benefits. Given that most of telehealth interventions are public projects and prices are regulated by payments systems such as DGR-rates, the entry of private providers, industry and insurance companies in the telehealth market will depend on the leadership of the national and/or regional authorities. Nonetheless, the microeconomic evidence on cost-effectiveness based upon scattered interventions does not enable assessment by the regional or national authorities of the fiscal sustainability of a large scale intervention.

Many of the different types of economic evaluations (cost-minimisation, cost-consequences, cost-effectiveness, cost-utility, and cost-benefit) of telehealth projects found in the literature do not inform about the viewpoint of the analysis and even when the study states that it responds to a “societal” perspective,\textsuperscript{45} there is no assessment of the fiscal impact or macroeconomic impact. An example of assessment of the fiscal impact of a telehealth intervention has been presented to assess the new reimbursement law for telehealth in California, Assembly Bill 415, focused on the Medi-Cal program\textsuperscript{46} which reimburses videoconferencing and some services for teleophthalmology, teledermatology and teleoptometry. Despite the fiscal/macroeconomic scope of this evaluation, the assessment is based on an extrapolation of results on costs and effects found in the literature to a comparable population eligible for home monitoring for heart failure and diabetes patients. The parametric assumptions taken in this study are a 42\% reduction in healthcare costs for heart failure due to telehealth, based on average from seven studies. Next, the estimation of Medi-Cal expenditure for heart failure patients, adjusted to heart failure costs (cost per capita five times higher than average) is applied a reduction of 42\% to estimate total potential savings in the Medi-Cal program, obtaining an upper limit of $8,600 per beneficiary annually. A similar reasoning is applied to diabetes beneficiaries of Medi-Cal, with a parametric assumption of a 9\% costs reduction due to telemonitoring based on research findings, which represents a savings of $939 per patient annually.

This treatment of “scalability” by simply extrapolating the impact found in small scale interventions, sometimes in controlled conditions such as in RCT, to larger populations facing different conditions may fail the assessment. For example, a small pilot on a support intervention for patients with heart failure in one centre based on randomized trial was evaluated and showed an impact of

\textsuperscript{45} Mistry (2012) Opus cit. reports about 12 studies in the literature from a societal perspective.
reduction in hospital readmission rate of 44%. This pilot study was followed by a large study on 1,653 patients in different centres and the evaluation study found no reduction in the risk of readmission or death from any cause with telehealth as compared with usual care. The correct methodology to link parameters found by using different microdata (providers and patients) in particular interventions to regional or national accounts such as healthcare expenditures in hospitalisation, reimbursement bills, productivity and or wages, even to GDP, is based on micro-macro models.

3.2 Micro-macro models on the impact of telehealth and telecare

On the one hand, the studies included in the systematic reviews on the impact of telehealth provide micro-level evidence either focused on impact for patients, mainly clinical, (disease-level approaches), or focused on efficiency in the use of resources of particular health providers (sub-sectoral approaches). On the other hand, the kind of evaluation at macro level of the impact of global ICT on the health sector has focused on the estimation of health cost functions to estimate and simulate economies of scale and scope, and efficiency analysis by using frontier analysis (DEA or SFA) by using providers (hospitals) at national level, and even in international comparisons (WHO and OECD). To move to the field of international comparisons, macroeconomic studies of the health sector have measured the impact of technology on health care expenditure as a residual accounting for supply-side factors or the non-demographic costs. This could be further linked to the macroeconomic impact of increasing life expectancy and health on economic growth.

Yet, there is no link between these two approaches, and the macro approach presented in these terms is too general to be applicable to assess the particular impact of telehealth and telecare as a part of ICT in health. A bridge between both approaches was presented under several micro-macro models in an EC funded project to assess the economic impact of ICT, defined at global level. In particular, the project presents a model which allows us to measure how the ICT intensity in public sector establishments (e.g. healthcare providers, government) affects the productivity impact of ICT investments in the private sector, or the impact of ICT in organisational management. However the type of data upon which these models have been constructed (19,000 firms across 13 countries with data from 1998 to 2008), which are required to estimate production functions at firm level, are not available at health care provider level. Besides, there is no consensus on the definition of a comparable aggregate output for healthcare providers to estimate a production function depending on non-ICT and ICT capital, labour, and other factors.

Nonetheless, the different eHealth benchmarking exercises carried out through surveys to hospitals, primary care centres and patients have resulted in different international comparable indicators which can serve as the micro-macro link needed to do a proper “scalation”. For example, the

---

hospital eHealth deployment index constructed from a hospital survey data\textsuperscript{56} can be used to further decompose effects of total ICT on health as measured from macroeconomic models based on growth accounting, DEA or panel data estimates. Also, these eHealth indexes can be used as explanatory variables in micro-macro models based on firm-level production functions (hospital costs functions in our case) of the type presented in the project on Economic Impact of ICT.

The considerations of spillovers to the overall economy can be applied after the basic impact on health has been measured through a “health equation”. Once an impact of telehealth deployment on a health indicator (e.g. life expectancy, mortality, health expenditure) is measured, this effect can be translated to the rest of the economy. By means of analogy, a basic model linking income and social indicators, such as life expectancy and education, to allow the analysis of policy shocks, has been developed by the World Bank as a basic macroeconomic accounting framework to project poverty.\textsuperscript{57} Such a model could be adapted to measure the impact of telehealth at macro level. More generally complete and dynamic frameworks linking income and social indicators, including the labour market, can be modelled by constructing a Social Accounting Matrix as the basis for a Micro-simulation Model or a Computable General Equilibrium Model.

4. CONCLUSIONS

This paper has reviewed the state of the art in assessment of telehealth with the current methodology which aims to integrate all stakeholders’ viewpoints. These new methodologies bridge traditional randomised trials methods employed in medico-economic impact analysis with more pragmatic methods which include qualitative assessments and also quantitative methods applicable under more realistic types of randomisation methods.

The trade-off between multidisciplinarity (HTA, MAST and GEMSA) and small-scale controlled experimentation (individual RCT applied under totally controlled conditions) translates to the trade-off between universality and causation achieved in the impact assessment. At the one end, the broadest and multidisciplinary methods such as MAST achieve to monitor a programme since Monitoring aims to be universal, treats the whole population, considers all kind of costs and benefits, and tracks implementation looking at what but not at why. At the other end, selected and narrow projects such as randomised control trials (individual randomisation) can succeed in measuring impact on the programme since Impact Evaluation is sample based and informs on the causes, but the assessment is not universal and then it is rarely scalable and transferable. This goes back to the trade-off and distinction between Monitoring and Operational Evaluation (M&OE) and Impact Evaluation (IE).\textsuperscript{58}

All in all, the result of the impact assessment, be it through M&OE or through IE, depends on the definition of the counterfactual as baseline scenario or do-nothing scenario. The status quo scenario plays a big role in the methodologies based on M&OE and multidisciplinary approaches. Therefore, in these kinds of impact assessment the definition of the counterfactual is not well adjusted to allow for the change of all of the factors which are changing with the telehealth policy intervention and also affect the result of this policy, but are not part of the impact of telehealth (confounders). This limitation is important for the measurement of impact in quantitative terms. To overcome that, sensibility analysis are performed and accompanied by qualitative descriptions of the context. In IE methods, the construction of a better baseline scenario to assess impact must discount these confounder effects. Nonetheless, the perfect counterfactual does not exist since we can only observe the realized outcome but not the potential. This is the fundamental problem of causal

\textsuperscript{56} Codagnone, C. and F. Lupianez-Villanueva, Benchmarking of eHealth Deployment in European Acute Hospitals: making sense of survey data and identifying the future research agenda 2011, JRC IPTS: Seville.

\textsuperscript{57} The original model was presented by the World Bank in 2001 as a Poverty Module (RMSM-X) and have been made available as a projection toolkit (PovStat 2.12).

\textsuperscript{58} The issue of the trade off between M&OE and IE was discussed in the Renewing Health Concertation Meeting held in Brussels on 29 March 2011.
inference and the optimal methods to assess telehealth impact and at the same time to do so in a scalable and transferable manner - that is, at macro-level - will always be subject to sensibility test of scenarios under different conditions. Hence, in our view, the current multidisciplinary impact assessment methods as well as more experimental and pragmatic methods such as WSD, are at least a necessary step towards achieving a reliable accuracy in impact assessment of telehealth applications.
Abstract

The challenges for healthcare systems in the European Union include demand side and supply side drivers. On the demand side, demographic changes due to ageing and increasing personal income are shaping growing expectations of healthcare services to increase quality and access, reduce disease burden, respond to emergency disease risks, and assist mobility and adaptation to the workplace. On the supply side, healthcare systems are under the pressure from limited budgets and the increasing complexity of healthcare provision which requires the management of investments in technology and interoperability of information flows alongside organizational changes.
As the Commission’s in-house science service, the Joint Research Centre’s mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new standards, methods and tools, and sharing and transferring its know-how to the Member States and international community.

Key policy areas include: environment and climate change; energy and transport; agriculture and food security; health and consumer protection; information society and digital agenda; safety and security including nuclear; all supported through a cross-cutting and multi-disciplinary approach.