Using modelling and simulation to conduct an economic evaluation of a complex technological innovation in healthcare*

Tiago Cravo Oliveira, James Barlow, Steffen Bayer

Abstract

Healthcare costs have been growing above gross domestic product in many developed countries for the last four decades. As demand for healthcare services keeps rising, this trend is likely to persist. Many expect that technological innovation will allow healthcare systems to improve quality while reducing costs. Paradoxically, technological change is considered the biggest contributor to the growth in healthcare expenditure. Economic evaluation methods have been used by policy and decision makers to identify those technologies that produce biggest value for money, yet their application is difficult under conditions of dynamic complexity. Using system dynamics, a modelling and simulation approach from the field of operational research, we model the use of remote video consultations (teleconsultations) in Portugal. The approach allows us to conduct an economic evaluation as well as gain an understanding of how the regional healthcare system changes after the introduction of the new technology. Using a single model, we assess multiple outcome measures from a variety of perspectives and under different scenarios. We also conduct cost-effectiveness and cost-minimisation analyses, using simulation results. The findings suggest that teleconsultations are a cost-effective way of reducing waiting times, as well as increasing the number of patients seen, although this comes at an increase in costs for providers.

Imperial College London, South Kensington Campus, London SW7 2AZ

Keywords: Economic evaluation, healthcare costs, technological innovation, teleconsultation

1 Introduction

Across the majority of developed countries, healthcare costs have been growing above gross domestic product for more than 40 years (Cutler, 1995; OECD, 2010; Reinhardt, Hussey and Anderson, 2004). While the financial sustainability of healthcare spending may become an issue in the future, the more urgent question is how long resources will continue to be allocated to healthcare before spending on other sectors of the economy has to be reduced. According to Chernew, Hirth and Cutler (2003), this may happen as early as 2039. In fact, it has been argued that a decade's real income gains have already been wiped out by healthcare cost growth in the United States (Auerbach and Kellermann, 2011). Neilson (2008) claims that high costs

*An initial version of this paper was presented at the 30th International Conference of the System Dynamics Society in St. Gallen, Switzerland, on July 22, 2012. The model has since been revised and the results presented here have not been previously published.

are not really a problem, since healthcare systems employ many workers who provide valued services, yet this view largely disregards the distributional challenges that will arise when, and if, non-healthcare spending has to be reduced, not to mention questions relating to equity and solidarity (Bodenheimer, 2005*b*).

Policy and decision makers expect that new technologies will allow healthcare systems to achieve more with less: increase quality while reducing costs. Yet, it is generally accepted that technological change has been the major driver behind the rise in healthcare costs (Bodenheimer, 2005*a*; Cutler, 1995; Newhouse, 1992). The majority of recent technological innovations are considered to be cost-increasing (Bosanquet, 2009; Cutler and McClellan, 2001), and while costs continue to soar, the marginal benefits are increasingly negligible (Skinner, Staiger and Fisher, 2006). A number of technological innovations do appear to have lower unit costs when compared to conventional procedures, yet their impact on utilisation (i.e. number of services) actually offsets any cost savings. New technologies can either substitute or complement conventional procedures. In the latter case, their implementation either leads to more patients, more units of care, or both. In certain situations, substitution and expansion occur simultaneously, obscuring the overall impact. Many new technologies also enable new models of care, with more people being able to access treatment closer to home, and with impacts in specialised care units downstream. These new models of care require changes to organisational roles and responsibilities, with unintended and unexpected consequences that can reach far beyond the initial site of implementation.

It is increasingly difficult to evaluate these technologies. Economic evaluation methods - chief among them cost-effectiviness analysis (CEA) - provide limited support for dealing with dynamic complexity and feedback effects. The choice of time frame can lead to an overestimation of cost-effectiveness in the short-run (Van de Wetering, Woertman and Adang, 2012). The choice of scale can have considerable impact on results, depending on whether it captures economies or diseconomies of scale. CEA requires choosing one unambiguous objective of the intervention and a clear dimension along which to assess effectiveness, which is not always possible with complex service innovations. Furthermore, one should assess final outcomes rather than intermediate effects, yet evidence on the final outcomes of new technologies is often low quality and based on small scale pilot projects. Many new technologies show promising results in terms of patient experience, service quality and productivity, but poor evidence regarding clinical effectiveness. New approaches to the evaluation of complex technological innovations are desirable.

2 Approach

Operational research's system dynamics modelling and simulation methodology provides a readily accessible way of minimising the impact of these limitations. We use a system dynamics approach to evaluate the consequences of a technological innovation on a regional care economy over time, using different perspectives, different scales, and multiple effectiveness measures, with the objective of determining its impact on utilisation and costs.

The technological innovation is teleconsultation: remote outpatient video-consultations connecting a consultant in a hospital with a general practitioner (GP) and a patient in primary care. The technology has been in use in Alentejo – a region in Portugal – since 1998. Alentejo represents a third of the Portuguese territory, but with a population density six times below the national average, urban dwellings are dispersed and primary care practices can be hundreds of kilometres away from hospitals. The average *alentejano* is older, poorer and less educated than the average Portuguese, which would indicate a bigger need for medical care, *ceteris paribus*. However, a poor public transportation network, large distances between providers, and a shortage of hospital consultants, all create barriers to access to specialised care. Teleconsultations were implemented as a means to see, diagnose and treat patients as close as possible to their homes and workplaces.

Even though the technology has been in use for more than a decade now, it has yet to reach significant spread and the evidence base for its benefits is limited and low quality (Davalos, 2009). While experience to date suggests considerable benefits for patients and efficiency gains for providers, the effects on the regional healthcare system as a whole are poorly understood. Much effort has gone into establishing the feasibility of the technology, as well as its clinical equivalence to face-to-face appointments. Little or no attention has been given to whether costs and consequences from small pilot projects would hold under bigger scale implementations. Furthermore, there is evidence that teleconsultations act both as a substitute and a complement for conventional outpatient care. The overall impact on utilisation is poorly understood and may actually offset the reported cost savings. Finally, providers bear the brunt of the investment, with patients and families reaping the majority of the benefits. This uneven distribution of costs and benefits means the choice of perspective used in any evaluation of teleconsultations will be important.

Economic evaluation methods, such as CEA, are difficult to apply to technologies like teleconsultation, which explains why so few evaluations have been published to date and why they focus on very specific outcomes (Agha, 2002; Johnston et al., 2004). This is unfortunate because it creates uncertainty for all parties involved. Patients and healthcare professionals who see the benefits of teleconsultation in everyday practice do not understand why spread is so slow. Policy makers are reticent to back a larger scale implementation of the technology based on limited and poor quality evidence. Technology providers are not sure whether to invest in products that may not reach practice.

With this research, we aim to comprehensively assess the costs and consequences of teleconsultations in Alentejo, from a number of perspectives, under different assumptions and scales, and over two periods of time, 5 and 10 years. System dynamics has been used in this context before. Smith and van Ackere (2002) have shown how system dynamics models can be integrated with micro-economic models to provide guidance to policy makers. At least two other studies have combined system dynamics with economic models of chronic care to assess time effects and distribution of costs and consequences (Bayer, Barlow and Curry, 2007; Homer, Hirsch and Milstein, 2007).

3 Methods

We conducted a literature review of real-time teleconsultations and organised semi-structured exploratory interviews with GPs, consultants and managers in Alentejo. The results of these two steps informed the development of the initial model. Previously published system dynamics models were used as references, whenever appropriate. While teleconsultations are used in many medical specialties, we restricted our analysis to three: dermatology, neurology and surgery. These were chosen so that the number of consultations provided was highest (i.e. higher sample sizes), and differences among the specialties might be modelled via parameters rather than whole different models. This way we may be able to compare different uses of the same technology, at a later stage in the research. We are now collecting data using both qualitative and quantitative methods: a dataset of first outpatient appointments in the region's main hospital, a discrete choice experiment of GP referrals, and a survey of patients.

3.1 Data

We have just recently been given a dataset of all the face-to-face first outpatient appointments in the three specialties under analysis from 2002-2011 featuring patient number, date of referral and date of appointment, patient age and gender, and primary care unit that referred. Combining this dataset with data on the evolution of supply throughout the same period (i.e. number of GPs and consultants), the supply of private care and the needs of the population, one can conduct an analysis similar to the one performed by Martin and Smith (1999), the results of which were used by Van Ackere and Smith (1999). Our analysis should provide important parameters for our model, most importantly estimates of the elasticities of demand (i.e. referrals and patient completion) with respect to waiting time and distance.

In the Portuguese National Health Service, GPs act as gatekeepers: a patient can only see an NHS consultant if he/she has been referred by a GP. With health information readily available online, the decision to refer is increasingly a shared one, with GPs and patients both having a say. With revealed data (i.e. the dataset described above) it is very difficult to disentangle the impact of changes in waiting time on patients and GPs individually. The use of stated choice methods can provide a way forward. A discrete choice experiment (DCE) is an attribute-based survey method based on the assumption that a good or service can be described by a set of characteristics or attributes and that the extent to which individuals value that good or service is determined by the nature and levels of the characteristics (Ryan et al., 2001). Our DCE presents GPs with a number of hypothetical patient cases (based on real cases) and they have to decide whether to refer, and with what priority. The attributes - identified in a literature review - are waiting time, distance, patient pressure and clinical need (the latter is inherent to each patient case). The method allows us to test and quantify the impact of these determinants on the choice of referral. The results will be used to improve the model.

Finally, many of the benefits of teleconsultations accrue to patients: less time off work, more leisure time, fewer travelling costs, and reduced waiting time, to name a few non-clinical benefits. In order to quantify, estimate and monetise these effects, we need information on means of transportation used and time taken. We must also understand whether patients come accompanied by family members or friends (frequently the case for elderly patients and children), so we can include these in the analysis. The patients' level of satisfaction with the consultation may also be an important measure of effectiveness. For all these reasons, we conducted a telephonic survey of patients who had attended a first outpatient appointment in the last semester of 2011, in the region's main hospital (100 patients in the conventional face-to-face pathway, and another 100 in the teleconsultation pathway).

4 The model

In the following paragraphs, we describe a highly aggregated and simplified version of the underlying simulation model used in the study (many variables are omitted for readability). In locations where teleconsultation is not available, patients referred by their GP are booked an outpatient face-to-face appointment, in one of the region's hospitals. They may need to travel

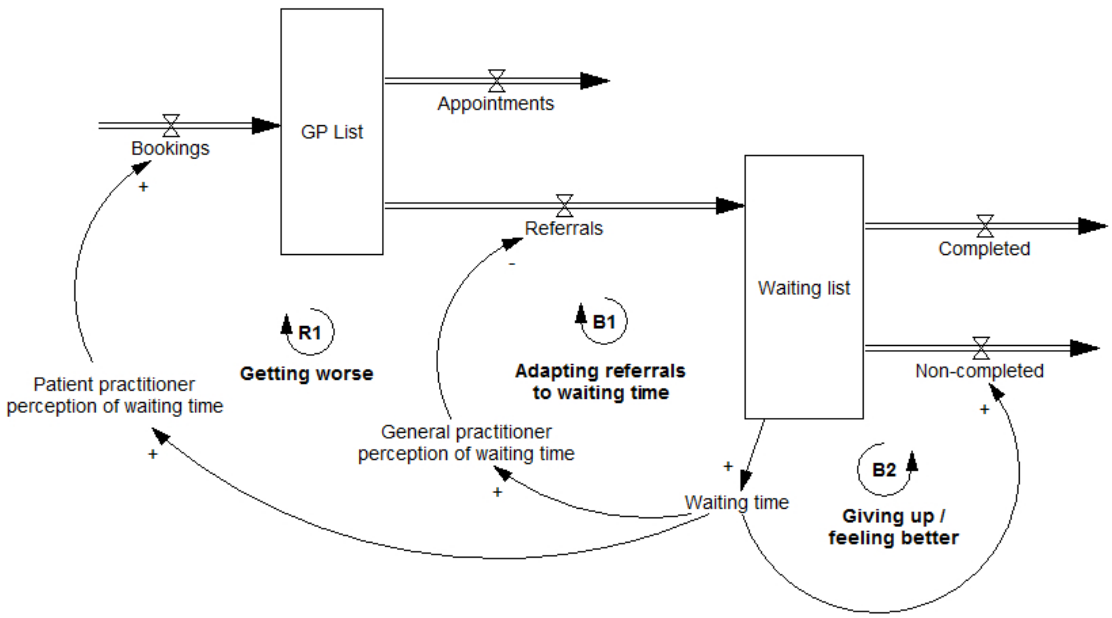

Fig. 1: Simplified patient flows.

more than a hundred kilometres each way, at great personal cost, both in time and money. Some, for example children and elderly, have to be accompanied by family or friends, while others are entitled to travel by ambulance, at the expense of providers. Due to a significant mismatch between the supply and distribution of consultants, and the size and needs of the population, waiting times reach, on average, four months. Figure 1 shows the flow of patients as they move towards an outpatient face-to-face appointment.

The routes are depicted by pipes (double lines) through which people (patients) move. Patients flow from infinite sources of people seeking care to infinite sinks of treated people (both shown as clouds), and their progress is regulated by rate of flow variables on the pipes (shown as valves). The rates of flow are often dependent on further variables, which may move in the same direction or in an opposite direction to the flow they affect (for example, the variable *Referral rate* affects the number of people being referred in such a way that an increase in the variable will cause an increase in the rate of flow, *ceteris paribus*). People flow into and out of different states, depicted by boxes, where they wait for further progress through the system. Feedback arises from patients' and GPs' perceptions of outpatient waiting time, affecting the number of referrals and the number of patients going to their primary care practice in the meantime. As Van Ackere and Smith (1999), we model these perceptions as stocks, which change as new information concerning the waiting time is acquired. In other words, the perception of waiting time is a smooth of the actual waiting time. This is an appropriate representation of how people adapt their expectations based on delayed information.

The introduction of teleconsultations in a number of primary care practices leads to a new pathway (see Figure 2). Primary care practices in the telemedicine network refer their patients to teleconsultations rather than to face-to-face appointments. Depending on the outcome of the teleconsultation, they may then be referred to a conventional face-to-face appointment, if

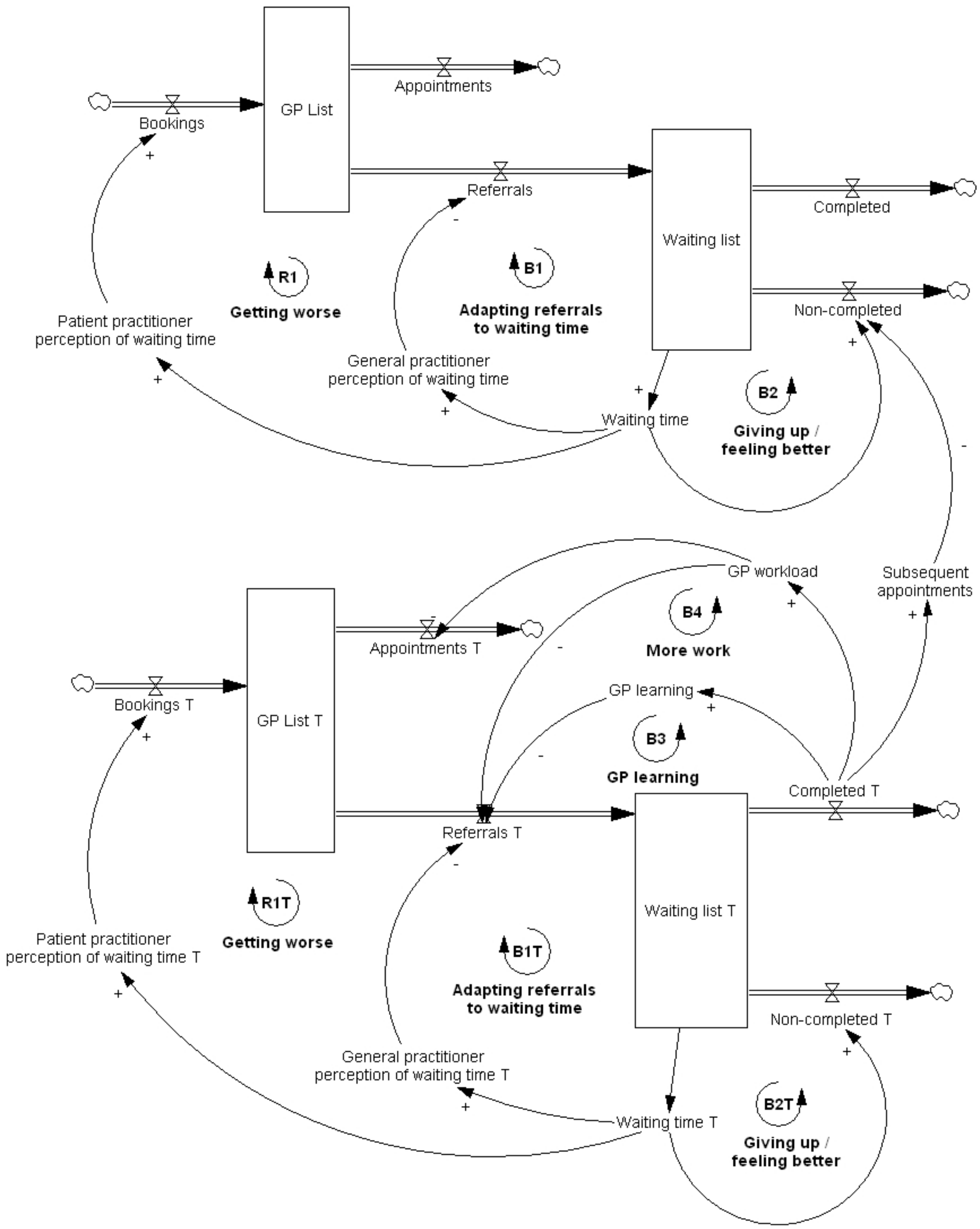

Fig. 2: System after introduction of teleconsultation.

needed. The teleconsultation is real-time with the presence of a GP (either the patient's or one from the same practice) and takes place at a specific time each week (for example, every Monday at 10am). Because there are two physicians involved in different locations, schedules are followed more strictly. There is less small talk during a teleconsultation and the clinical examination is facilitated by the GP. Thus, teleconsultations are quicker than face-to-face appointments. Patient data is readily available via an electronic patient record system and the consultant can request diagnostic procedures as in a conventional face-to-face appointment. Once a diagnosis is made, treatment can be prescribed either by the GP or by the consultant (certain medications must be prescribed by a consultant, e.g. neurological drugs). Prescriptions written by consultants are mailed directly to patients. If a remote diagnosis is not possible, then a face-to-face consultation is scheduled. Administrative staff manage the booking schedules and operate the videoconferencing equipment.

The fact that GPs are present during teleconsultations leads to two new feedback loops. On the one hand, GPs gain experience from participating in outpatient appointments, they gain knowledge about specialist care, increasing their ability to diagnose and treat patients in primary care. This leads to a reduction in the number of patients that have to be referred. On the other hand, GPs are taking on new work, which reduces the time available for primary care appointments. This may eventually lead to one of two things: higher human resource costs, or permanently lower capacity in primary care. In our model, we assume lower capacity in primary care, and thus expect primary care waiting lists in practices using teleconsultation to increase.

While the two pathways (face-to-face and teleconsultation) are mostly independent of one another in primary care, they share an essential secondary care component: consultants. The consultants that perform face-to-face appointments are the same ones that do teleconsultations. This effectively means that teleconsultation patients compete with face-to-face patients for consultants' time. The more time consultants spend doing teleconsultations the less time they have for face-to-face appointments, and vice-versa (we assume there is no possibility of extra hours, trade-offs have to be made). Managing the allocation of time to the two pathways is far from easy, and may even lead to inequities in access between practices with teleconsultation and those without.

4.1 Parameters

The model allows us to explore the impact of teleconsultations on a number of indicators (e.g. waiting times, non-completed appointments, costs, etc.), from a variety of perspectives (e.g. societal, patient, national health service, individual practices, etc.), under different scenarios (e.g. scales and policies). Calibration of the model requires the estimation of a range of parameters (see Table 1). The effect of waiting times on other variables was estimated based on a study determining the elasticity of demand and the elasticity of private insurance, both with respect to waiting time (Van Ackere and Smith, 1999). Figure 4 illustrates how the elasticity of demand changes with respect to waiting time. The values chosen for subsequent appointments, non-completion and learning, were based on interviews in Alentejo, and published studies (Hersh et al., 2006; Lamminen et al., 2001; Loane et al., 2000; Oakley et al., 2000; Sicotte et al., 2004; Jaatinen et al., 2002; Nordal et al., 2001; Forrest et al., 2007).

To operationalise the effect of GP learning, we use an s-curve. Based on interviews, it would take approximately 6 months of teleconsultations to reduce referrals by 50 percent. We calibrate the curve so that after 6 months of teleconsultations, GPs using the technology reduce

Tab. 1: General model parameters.

Parameter (unit)	**Value**	**Source**
Initial primary care waiting time (week)	1	Interviews
Initial outpatient waiting time (week)	20	(HESE, 2011)
Dermatologists (physician)	2	(Natario and Amaral, 2011)
Need (consultations per 1000 people)	50	(Natario and Amaral, 2011)
Duration of primary care appointment (hour)	0.25	(Natario and Amaral, 2011)
Duration of dermatology consultation	0.2	Interviews
Duration of dermatology teleconsultation	0.1	Interviews
Initial referral rate (per 1000 GP appointments)	22	(INE, 2010)
Initial share of population with private insurance	17%	(Natario and Amaral, 2011)
Time to perceive waiting time (week)	52	(Van Ackere and Smith, 1999)
Elasticity of private insurance	0.3	(Van Ackere and Smith, 1999)
Non-completion share	25%	(Forrest et al., 2007)
Cost of each video-conferencing platform (€)	7,200	Interviews
Cost of maintaining equipment (€)	72	Estimated
Patients travelling by car (face-to-face)	80%	Patient survey
Patients travelling by bus (face-to-face)	8%	Patient survey
Patients travelling by taxi (face-to-face)	1%	Patient survey
Patients travelling by ambulance (face-to-face)	11%	Patient survey
Patients travelling by car (teleconsultation)	57%	Patient survey
Patients travelling by bus (teleconsultation)	2%	Patient survey
Patients travelling by taxi (teleconsultation)	0%	Patient survey
Patients travelling by ambulance (teleconsultation)	2%	Patient survey
Patients walking to teleconsultation	39%	Patient survey
Distance to hospital (km)	47	Patient survey
Distance to primary care practice (km)	6	Patient survey
Average hourly wage (€/hour)	8.3	(INE, 2010)
Share of population working	17%	(INE, 2010)
Time to attend face-to-face appointment (hour)	3	Patient survey
Time to attend teleconsultation (hour)	1.5	Patient survey
Co-payment, face-to-face (€)	7.5	(Diário da República, 2011)
Co-payment, teleconsultation (€)	3	(Diário da República, 2011)
Reimbursement, first appointments (€)	121.07	(ACSS, 2012)
Reimbursement, subsequent appointments (€)	110.07	(ACSS, 2012)

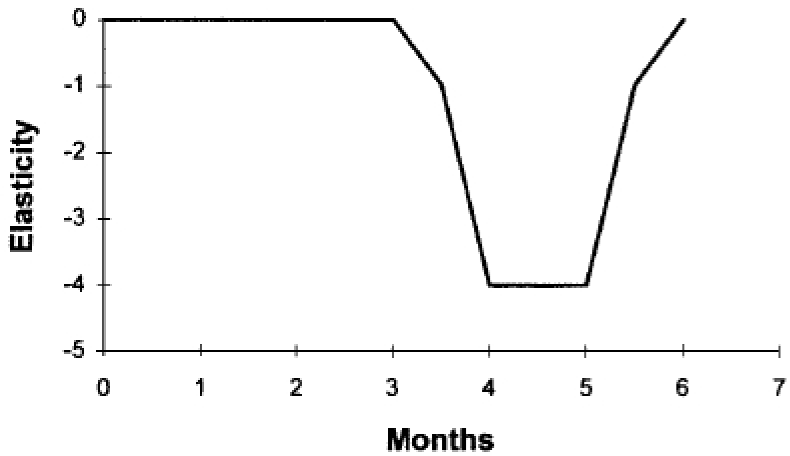

Fig. 3: Elasticity of demand with respect to waiting time (Van Ackere and Smith, 1999).

their referrals by 50 percent. We assume experience acquired through teleconsultations is not transferable to other GPs, hence only GPs participating in teleconsultations are able to reduce their referrals. We consider an s-curve to be appropriate because there is an initial period of slow improvement, as GPs adapt to secondary care methods and ways of working; as time passes, there is rapid improvement in their ability to diagnose and treat a growing number of patients themselves without having to refer; after a longer time period, GPs reach the limit of what they are able to do in primary care before having to refer patients onwards. This rationale can be challenged in the future, once validation work is under way. We assume there are no negative productivity effects in the early stages of learning (e.g. the duration of the teleconsultation is the same irrespective of the stage of learning; this can be challenged in the near future).

We have just recently been given access to the production dataset and have not yet analysed all the information contained in the detail. Thus, we assess the use of telemedicine in a hypothetical primary care practice serving a population of 10,500, and employing 7 GPs (in accordance with guidelines from the Portuguese NHS that there should be one GP for every 1,500 residents). All 7 GPs participate in teleconsultations. We exclude from the analysis the possibility that patients will use of primary care practices when waiting time for outpatient appointments increases (loops R1 and R1T), and the costs of administrative staff. The reason is we have limited data and understanding of these issues and so it is premature to include them in the simulations.

After taking into account the estimated parameters and the constraints imposed by the logical and arithmetical relationships in the model, one important parameter remained to be determined: consultant work hours. Outpatient appointments are only one of the activities that consultants have to perform. Furthermore, if the analysis focuses on only a specific number of practices, then we have to make a judgement concerning the amount of time that consultants spend with patients from those practices, relative to all their outpatients. Consultant work hours were determined so that the system would be in a steady-state before the introduction of teleconsultations. In other words, the number of patients getting referred is the same as the number of patients being seen, and so the initial waiting time of 20 weeks remains constant throughout the simulation period. We also assume that, after teleconsultation is introduced, consultants allocate the same amount of time to teleconsultations as they did before for face-to-face appointments.

5 Early findings

Having developed and parameterised the model, we can now run a number of simulation experiments to assess the impact of teleconsultations on provider, patient and societal costs, as well as different measures of non-clinical effectiveness (we are currently working towards the inclusion of clinical effects in the analysis). We assume that after 10 years the video-conferencing equipment has to be replaced. The care system is composed of only one hospital and one primary care practice (in the following section we will include more practices to assess the impact of scale on the results). Here we report on findings for dermatology (work on the other two specialties is ongoing). We conducted three simulation runs to gain an initial understanding of how the effects of teleconsultations reported in the literature and interviews impacted on costs and consequences:

A: Steady-state system before introduction of teleconsultation.

B: System after introduction of teleconsultation: 10 percent of teleconsultations result in a subsequent face-to-face appointment; reduction of 20 percent in non-completion of appointments.

C: Full model: Run B plus GPs reduce their referrals by half, six months after starting to use teleconsultations.

The difference between runs B and C is the inclusion of the general practitioner learning effect. As Figure 4 illustrates, the inclusion of this effect can have a significant impact on the results of the evaluation. Teleconsultations take half the time of a face-to-face appointment (see Table 1), effectively increasing service capacity and reducing the waiting time. This reduction in turn leads to an increase in the number of referrals. As the perception of waiting time is based on delayed information concerning the actual waiting time, in the initial months of using teleconsultations there are oscillations. As the actual waiting time and the perceived waiting time finally align, the system seems to reach equilibrium around 120 days in run B (see Figure 4). However, this is not the case for run C, which exhibits an impressive decline in the waiting time. This illustrates how important the inclusion of all consequences can be: general practitioner learning has an astounding impact on results. Missing important information can have significant implications when evaluating the costs and consequences of teleconsultations.

5.1 Utilisation and costs

Table 2 shows the impact of the technology on effectiveness and cost measures, for a 10 year time frame and for runs A (face-to-face) and C (teleconsultation). As seen in Figures 5 and 6, teleconsultations can have a significant impact on waiting time. But there are also significant changes in utilisation rates. The referral rate increases from 22 referrals per 1000 primary care appointments to 41 referrals per 1000 primary care appointments (this would be even higher if it were not for general practitioner learning). A total of 766 more patients get a dermatology

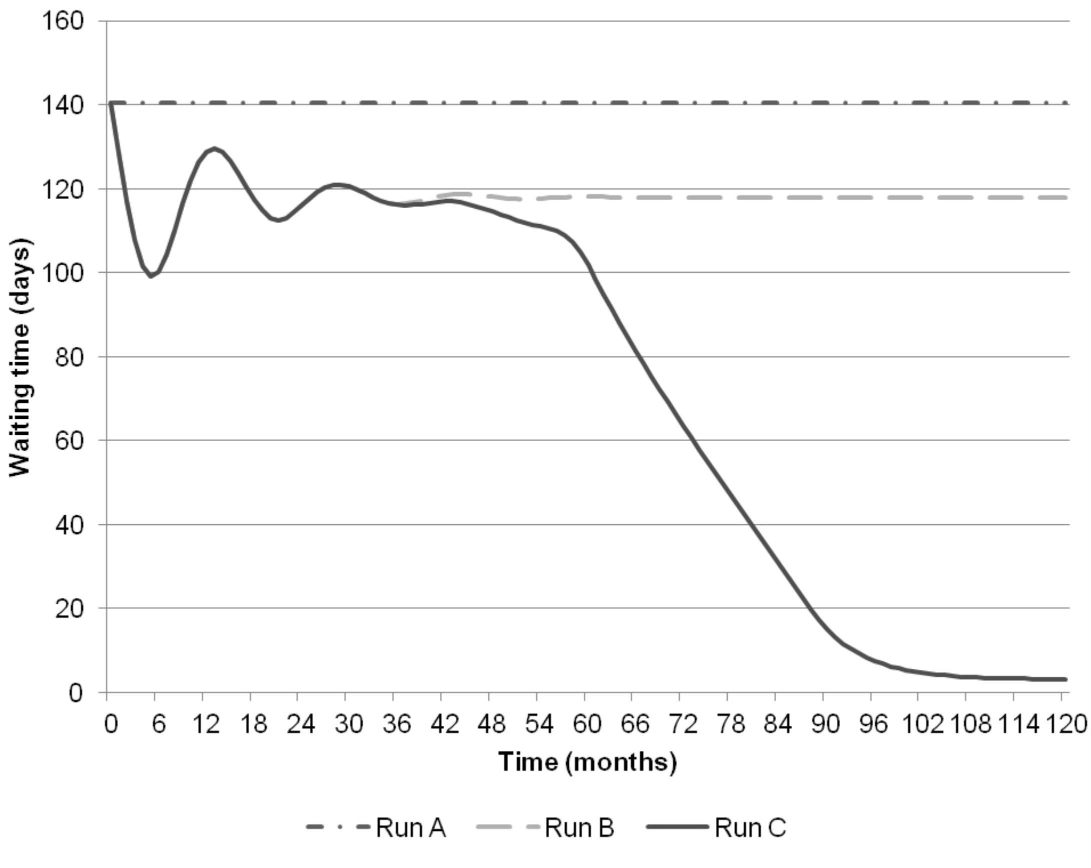

Fig. 4: Waiting times for outpatient appointments for a 10-year simulation.

Tab. 2: Impact measures for a 10-year simulation of full model.

Impact measures (unit)	**Face-to-face**	**Teleconsultation**	△
Outpatient waiting time (days)	140	10	-130
Primary care waiting time (days)	7	13	6
Referral rate (per 1000 GP appointments)	22	41	19
Share of population with private insurance	17%	8%	9%
Primary care production (patients)	315,013	314,491	522
Secondary care production (patients)	529	1,295	766
Non-completion (patients)	176	68	108
Subsequent face-to-face (patients)	0	129	129
Avoided referrals (patients)	0	770	770
Total costs to society (€)	2,308,048	2,419,703	111,655
Total costs to NHS (€)	1,350,600	1,466,811	116,211
Total costs to patients (€)	957,047	952,858	-4,189

consultation after the introduction of teleconsultation, 108 patients attend their consultation when they would not before, 129 patients get a subsequent face-to-face appointment after a teleconsultation, and 770 patients get their problem taken care of in primary care without being referred to a teleconsultation. On the other hand, because general practitioners have to be present in teleconsultations, they see 522 fewer patients in primary care. We can now determine different measures of cost-outcomes ratios.

We can calculate a cost-effectiveness ratio using outpatient waiting time as the measure of effectiveness. In the full model, the waiting time has improved by 130 days and total NHS spending has increased by €116,211. Thus, each day each of the 1,295 patients did not have to wait cost the NHS €0.69. For patients, teleconsultations are a dominant alternative to face-to-face appointments, because they reduce both costs and waiting times. The societal cost of each patient's avoided waiting day is €0.66.

We can also conduct a cost minimisation analysis, in which we match utilisation in both pathways (teleconsultation and conventional face-to-face). So, for example, for the full model, general practitioners would need to see 522 more patients in the teleconsultation pathway. In the conventional pathway, consultants would have to see 766 more patients plus 770 who would have been referred if teleconsultation was not available. The total societal saving is €65,240, indicating that teleconsultation is a less costly way to provide care at this level of utilisation. This does not seem much money over 10 years, but remember this is for only one specialty and one practice.

5.2 Scale effects

We now analyse a system with four primary care practices. Population and number of general practitioners are increased accordingly. The number of consultants is kept constant at two, yet the amount of time allocated is increased so that, as earlier, the number of patients referred equals the number of patients seen, before the introduction of teleconsultation. We conducted five more runs:

D: Steady-state system before teleconsultation.

E: System after introduction of teleconsultation in one practice.

F: System after introduction of teleconsultation in two practices.

G: System after introduction of teleconsultation in three practices.

H: System after the introduction of teleconsultation in all four practices.

We use the full model (first used in run C) for runs D through H. When a practice implements the technology, 10 percent of teleconsultations in that practice will lead to a subsequent face-to-face appointment. We assume (in accordance with interview data) that these patients do not get added to the face-to-face outpatient waiting list. Instead, they are booked immediately and reduce the number of face-to-face appointments consultants can perform for patients from the respective waiting list. In effect, patients from practices without teleconsultation will experience reduced capacity, and the reductions will be bigger as the number of practices using teleconsultation increases. This effect is clearly visible in Table 3 and would be missed in studies which do not take into account scale of implementation.

In run G, for example, 3 practices have implemented teleconsultations, leaving only one practice in the face-to-face pathway. The number of patients needing a subsequent face-to-face appointment is now 409. This number represents 62 percent of consultants' time allocated to the face-to-face practice. The result is an increase in the waiting time for face-to-face patients to more than 8 years. In reality, this would never happen, as healthcare professionals would take steps to avoid it (e.g. increase the time allocated to face-to-face appointments or hire more consultants, etc.). However, what this set of runs (D through H) shows is that teleconsultations can lead to significant reductions in the amount of time consultants have for patients from practices without teleconsultation. One practice's choice to implement the technology can have significant effects elsewhere in the system. To avoid this problem, healthcare managers have to be careful when allocating consultants' time to different practices. This is not easy, but it is essential if we are to maintain equity between practices.

The effect of scale on savings is also clearer if we conduct a cost-minimisation analysis, as earlier. The potential societal savings if all four practices adopt teleconsultation (run H) are €434,967, with providers saving approximately €283,428. However, one has to be careful when considering these numbers. They illustrate how much cheaper teleconsultation is compared to seeing all these patients using regular existing methods. In other words, if we were to see the same number of patients using conventional face-to-face consultations, this would be more expensive. However, in real terms, the NHS budget is actually inflated by between €103,876 (run E) and €443,528 (run H). Even so, considering the latter number is for four practices, over 10 years, teleconsultations can probably be considered a cost neutral innovation (i.e. do not change costs but change effectiveness).

6 Next steps

We are currently working on simulation runs in which the capacity of teleconsultations is limited to meet some measure of demand. We can also limit the amount of patients seen in the

Tab. 3: Impact of scale on cost and effectiveness measures for a 10-year simulation.

Runs	**Cost and effectiveness measures** Face-to-face waiting time (days)	Primary care production (patients	Secondary care production (patients)	Non-completion (patients	Subsequent face-to-face (patients)	Share of face-to-face spent on subsequent	Avoided referrals (patients)	Total costs to the NHS (€'000)	Total costs to patients (€'000)	Total costs to society (€'000)
D	140	1,260,052	2,117	706	0	0	0	5,402	3,828	9,232
E	143	1,259,442	2,785	565	129	8%	770	5,506	3,821	9,328
△*	*3*	*-611*	*668*	*-141*	*129*	*8%*	*770*	*104*	*-7*	*96*
F	149	1,259,009	3,456	424	259	23%	1,541	5,603	3,815	9,419
△*	*9*	*-1,043*	*1,340*	*-281*	*259*	*23%*	*1,541*	*201*	*-13*	*187*
G	2,458	1,258,328	4,127	284	389	62%	2,312	5,700	3,808	9,508
△*	*2,318*	*-1,726*	*2,011*	*-422*	*389*	*62%*	*2,312*	*298*	*-20*	*276*
H	NA	1,257,965	5,188	273	519	NA	3,083	5,846	3,812	9,658
△*	*NA*	*-2,087*	*3,071*	*-433*	*519*	*NA*	*3,083*	*444*	*-16*	*424*

* Differences to run D (e.g. value in run E minus value in run D)

teleconsultation pathway so that it equals the number of patients seen in the face-to-face pathway. In essence, these exercises and the runs we did up to now (A through H), are evaluations of the results starting from different assumptions concerning the number of desired consultations and the policies to achieve them. One important thing to bear in mind is that the initial investment in video-conferencing equipment creates an incentive to see more people, i.e. to dilute the fixed costs among a bigger number of patients. This effectively means that the bigger the number of teleconsultations performed, the bigger the savings compared to seeing those patients in the face-to-face care pathway. Our intuition is that the lower the number of consultations (i.e. due to policies introducing ceilings on production), the lower the hypothetical savings.

We are presently exploring and questioning some of the results in more detail. One result of interest is the effect of scale, i.e. the impact of one practice's decision to use teleconsultations on another practice's service. Questions which relate to this issue and remain unanswered include: who decides how much consultant time is spent on face-to-face appointments and teleconsultations? What influences the decision? Can managers and consultants measure how much time is spent seeing patients from each practice and can they create mechanisms to effectively change it? We are also experimenting with the possibility of making non-completion rates dependent on waiting time. We have data on missed appointments and waiting times, so we can calculate the elasticity of the non-completion rate with respect to waiting time (similar to the mechanism used in loop B1).

Another area which requires further work is the effect of general practitioner ability on referral rates. This is an important component of the model, with considerable impact on waiting time and savings. We are working towards obtaining validated estimates of the parameters involved (e.g. the amount of time or work needed to achieve a reduction in referrals and the magnitude of that reduction). One final issue is the impact of teleconsultations on clinical effectiveness. With waiting time decreasing by as much as 130 days, one would expect there to be positive clinical effects. Although a clinical trial is not feasible, we may be able to achieve an understanding of the effect by using aggregate data on the severity of the conditions diagnosed. Finally, data from our DCE should enable us to model the effect of waiting time and distance on referral rates. The results can also be compared to previously published elasticities (e.g. Martin and Smith (1999)).

We do not yet have complete data on waiting times, waiting lists and production in Alentejo. As such, it is not possible at this time to assess historical fit (in part, this is why our simulation experiments are based on hypothetical numbers, even though these are grounded in real available data). We have nonetheless performed a number of extreme conditions tests, reproducing the same tests performed by Van Ackere and Smith (1999) with similar results. We are thus confident that the model is not flawed in any significant way. We are presently unable to perform further tests of validity, as the model is still under development. We will also be conducting sensitivity analyses once more data is available. We find it is premature to conduct a sensitivity analysis at this point.

7 Conclusion

The findings presented here suggest that teleconsultations are a cost-effective way to reduce waiting time for outpatient appointments due to their combined effect on referral rates and productivity. They also seem to be a less costly way of increasing the number of patients seen, when compared to increasing face-to-face capacity. However, there are organisational challenges when both teleconsultations and face-to-face appointments are used in different primary care practices. We have shown how sensitive the results of the analysis are to the inclusion or omission of the different effects recognized in the literature (runs A through C), as well as how scale can affect the regional healthcare system as a whole (runs D through H). Future work will focus on establishing a measure of the validity of the model and its results, as well as conducting sensitivity analyses.

This work illustrates how modelling and simulation can be a complement to usual economic evaluation methods in the assessment of complex technological innovations, such as teleconsultations, without loss of information. We have shown that a single model allows for a variety of analyses (including cost-effectiveness and cost-minimisation) while still producing insights into how the system changes following the introduction of the new technology. Using only one model, it is possible to assess multiple outcome measures (e.g. waiting times, non-completed appointments, costs, etc.), from a variety of perspectives (e.g. societal, patient, national health service, individual practices, etc.), under different scenarios (e.g. scales and organisational models). The completeness of this approach should be of value to healthcare policy makers and analysts.

References

ACSS. 2012. "Contrato-Programa 2012: UOFC - Metodologia para definição de preços e fixação de objectivos.".

Agha, Z. 2002. "Cost effectiveness of telemedicine for the delivery of outpatient pulmonary care to a rural population." *Telemedicine and E-Health* 8(3):281.

Auerbach, David I. and Arthur L. Kellermann. 2011. "A Decade Of Health Care Cost Growth Has Wiped Out Real Income Gains For An Average US Family." *Health Affairs* 30(9):1630–1636.

Bayer, Steffen, James Barlow and Richard Curry. 2007. "Assessing the impact of a care innovation: telecare." *System Dynamics Review* 23(1):61–80.

Bodenheimer, Thomas. 2005*a*. "High and Rising Health Care Costs. Part 1: Seeking an Explanation." *Annals of Internal Medicine* 142(10):847–854.

Bodenheimer, Thomas. 2005*b*. "High and Rising Health Care Costs. Part 2: Technologic Innovation." *Annals of Internal Medicine* 142(11):932–937.

Bosanquet, Nick. 2009. *Technology: scientific force or power force*. Oxford University Press chapter 3, pp. 43–52.

Chernew, Michael E., Richard A. Hirth and David M. Cutler. 2003. "Increased Spending On Health Care: How Much Can The United States Afford?" *Health Affairs* 22(4):15–25.

Cutler, David M. 1995. "Technology, Health Costs, and the NIH.".

Cutler, David M. and Mark McClellan. 2001. "Is Technological Change In Medicine Worth It?" *Health Affairs* 20(5):11–29.

Davalos, M E. 2009. "Economic evaluation of telemedicine: review of the literature and research guidelines for benefit-cost analysis." *Telemedicine and e-health* 15(10):933–948.

Forrest, Christopher B., Efrat Shadmi, Paul A. Nutting and Barbara Starfield. 2007. "Specialty Referral Completion Among Primary Care Patients: Results From the ASPN Referral Study." *Ann Fam Med* 5(4):361–367.

Hersh, William R, David H Hickam, Susan M Severance, Tracy L Dana, Kathryn Pyle Krages and Mark Helfand. 2006. "Diagnosis, access and outcomes: update of a systematic review of telemedicine services." *J Telemed Telecare* 12(S2):3–31.

HESE. 2011. "Lista de Espera de Consultas - Setembro 2011.".

Homer, Jack, Gary Hirsch and Bobby Milstein. 2007. "Chronic illness in a complex health economy: the perils and promises of downstream and upstream reforms." *System Dynamics Review* 23(2-3):313–343.

INE. 2010. Anuário estatístico da região Alentejo 2010. Technical report Instituto Nacional de Estatística.

Jaatinen, Pekka T, Pertti Aarnio, Jouko Remes, Jorma Hannukainen and Tarja Koymari-Seilonen. 2002. "Teleconsultation as a replacement for referral to an outpatient clinic." *J Telemed Telecare* 8(2):102–106.

Johnston, K, C Kennedy, I Murdoch, P Taylor and C Cook. 2004. "The cost-effectiveness of technology transfer using telemedicine." *Health Policy and Planning* 19(5):302–309.

Lamminen, Heikki, Johanna Lamminen, Keijo Ruohonen and Hannu Uusitalo. 2001. "A cost study of teleconsultation for primary-care ophthalmology and dermatology." *J Telemed Telecare* 7(3):167–173.

Loane, M. A., S. E. Bloomer, R. Corbett, D. J. Eedy, N. Hicks, H. E. Lotery, C. Mathews, J. Paisley, K. Steele and R. Wootton. 2000. "A comparison of real-time and store-and-forward teledermatology: a cost-benefit study." *British Journal of Dermatology* 143(6):1241–1247.

Martin, Stephen and Peter C. Smith. 1999. "Rationing by waiting lists: an empirical investigation." *Journal of Public Economics* 71(1):141–164.

Natario, A and JC Amaral. 2011. Actuais e futuras necessidades previsionais de médicos (SNS). Technical report Administração Central do Sistema de Saúde (ACSS).

Neilson, Eric G. 2008. "Baumol's Curse on Medicine." *Journal of the American Society of Nephrology* 19(6):1049–1053.

Newhouse, Joseph P. 1992. "Medical Care Costs: How Much Welfare Loss?" *The Journal of Economic Perspectives* 6(3):3–21.

Nordal, E J, D Moseng, B Kvammen and M-L Lochen. 2001. "A comparative study of teleconsultations versus face-to-face consultations." *J Telemed Telecare* 7(5):257–265.

Oakley, A M M, P Kerr, M Duffill, M Rademaker, P Fleischl and N Bradford. 2000. "Patient cost-benefits of realtime teledermatology - a comparison of data from Northern Ireland and New Zealand." *J Telemed Telecare* 6(2):97–101.

OECD. 2010. *Health at a Glance: Europe 2010*. OECD Publishing.

Portaria nº 306-A/2011 de 20 de Dezembro. 2011.

Reinhardt, Uwe E., Peter S. Hussey and Gerard F. Anderson. 2004. "U.S. Health Care Spending In An International Context." *Health Affairs* 23(3):10–25.

Ryan, M, A Bate, C J Eastmond and A Ludbrook. 2001. "Use of discrete choice experiments to elicit preferences." *Quality in Health Care* 10(suppl 1):i55–i60.

Sicotte, Claude, Pascale Lehoux, Nicolaas Van Doesburg, Godefroy Cardinal and Yves Leblanc. 2004. "A cost-effectiveness analysis of interactive paediatric telecardiology." *J Telemed Telecare* 10(2):78–83.

Skinner, Jonathan S., Douglas O. Staiger and Elliott S. Fisher. 2006. "Is Technological Change In Medicine Always Worth It? The Case Of Acute Myocardial Infarction." *Health Affairs* 25(2):w34–w47.

Smith, Peter C. and Ann van Ackere. 2002. "A note on the integration of system dynamics and economic models." *Journal of Economic Dynamics and Control* 26(1):1–10.

Van Ackere, Ann and Peter C. Smith. 1999. "Towards a macro model of National Health Service waiting lists." *System Dynamics Review* 15(3):225–252.

Van de Wetering, Gijs, Willem H. Woertman and Eddy M. M. Adang. 2012. "A model to correct for short-run inefficiencies in economic evaluations in healthcare." *Health Economics* 21(3):270–281.