



Avaliação de desempenho das unidades de saúde para a melhoria dos cuidados de saúde primários

Autor(es): Gouveia, M. C.; Dias, L. C.; Antunes, C. H.; Mota, M. A.; Duarte, E. M.; Tenreiro, E. M.

Publicado por: Imprensa da Universidade de Coimbra

URL persistente: URI:<http://hdl.handle.net/10316.2/35940>

DOI: DOI:http://dx.doi.org/10.14195/978-989-26-0738-2_18

Accessed : 17-Jun-2017 10:38:18

A navegação consulta e descarregamento dos títulos inseridos nas Bibliotecas Digitais UC Digitalis, UC Pombalina e UC Impactum, pressupõem a aceitação plena e sem reservas dos Termos e Condições de Uso destas Bibliotecas Digitais, disponíveis em <https://digitalis.uc.pt/pt-pt/termos>.

Conforme exposto nos referidos Termos e Condições de Uso, o descarregamento de títulos de acesso restrito requer uma licença válida de autorização devendo o utilizador aceder ao(s) documento(s) a partir de um endereço de IP da instituição detentora da supramencionada licença.

Ao utilizador é apenas permitido o descarregamento para uso pessoal, pelo que o emprego do(s) título(s) descarregado(s) para outro fim, designadamente comercial, carece de autorização do respetivo autor ou editor da obra.

Na medida em que todas as obras da UC Digitalis se encontram protegidas pelo Código do Direito de Autor e Direitos Conexos e demais legislação aplicável, toda a cópia, parcial ou total, deste documento, nos casos em que é legalmente admitida, deverá conter ou fazer-se acompanhar por este aviso.



INVESTIGAÇÃO OPERACIONAL EM AÇÃO

CASOS DE APLICAÇÃO

RUI CARVALHO OLIVEIRA
JOSÉ SOEIRO FERREIRA
(EDITORES)

IMPRESA DA UNIVERSIDADE DE COIMBRA
COIMBRA UNIVERSITY PRESS

CASO 18

AVALIAÇÃO DE DESEMPENHO DAS UNIDADES DE SAÚDE PARA A MELHORIA DOS CUIDADOS DE SAÚDE PRIMÁRIOS

M.C. Gouveia

IPC - ISCA Coimbra
INESC Coimbra
mgouveia@iscac.pt

L.C. Dias

Faculdade de Economia, Universidade de Coimbra
INESC Coimbra
lmc dias@fe.uc.pt

C.H. Antunes

DEEC-FCT Universidade de Coimbra
INESC Coimbra
cantunes@inescc.pt

M.A. Mota

ACES Baixo Mondego I
mariaaugustamota@gmail.com

E.M. Duarte

ACES Baixo Mondego I
emrduarte@gmail.com

E.M. Tenreiro

ACES Baixo Mondego I
Eduardo.tenreiro@gmail.com

RESUMO

Este artigo relata um estudo de avaliação de desempenho de unidades de saúde prestadoras de cuidados de saúde primários realizado com o envolvimento dos decisores. Foram utilizadas metodologias, algumas desenvolvidas para este caso em particular, que exploram as ligações entre a Data Envelopment Analysis (DEA) e Análise de Decisão Multicritério (Multi-Criteria Decision Analysis - MCDA), para incorporar preferências de gestão na avaliação de desempenho de doze unidades com dados relativos ao ano de 2010. Neste trabalho são apresentados dois modelos distintos, resultantes de discussões com um grupo de decisores, que combinam aspectos de avaliação como o acesso, os recursos utilizados, os serviços prestados e os resultados alcançados nos Cuidados de Saúde Primários. É utilizada uma variante do modelo DEA aditivo em duas fases para a avaliação de eficiência. Esta análise contribui para a identificação das melhores práticas, das fontes de ineficiência, bem como das diferenças em relação às melhores práticas. O interesse da abordagem reside não só na contribuição para conceber medidas corretivas adequadas baseadas nesses resultados, mas também na vertente de proporcionar informação relevante para a tomada de decisões mais fundamentadas sobre as metas dos indicadores de desempenho no futuro.

PALAVRAS-CHAVE

Data Envelopment Analysis; Análise de eficiência; Análise Multicritério; Cuidados de saúde primários.

1. Introdução

Os Cuidados de Saúde Primários (CSP) são considerados de crucial importância no sistema de saúde (SNS) português. A experiência internacional demonstra que os modelos de organização e gestão dos CSP, bem como as estratégias de financiamento adotadas e a forma de articulação com os restantes níveis de cuidados, podem potenciar ou, pelo contrário, comprometer o cumprimento dos princípios basilares dos sistemas de saúde, bem como das metas definidas pela política de saúde ao nível da equidade, acesso, eficiência e qualidade dos cuidados.

A reforma dos CSP, iniciada em 2005, tem como propósito melhorar o acesso aos cuidados de saúde e a qualidade dos serviços prestados. Ao mesmo tempo, visa ganhos de eficiência, porquanto a dimensão do SNS é de tal ordem que práticas inadequadas podem causar enormes desperdícios. O impacto desta reforma está agora a ser testemunhado, o que justifica a necessidade de avaliar os seus resultados de modo a identificar as melhores práticas nos CSP.

O objetivo do presente trabalho é avaliar a eficiência das unidades de CSP monitorizadas pelo Agrupamento de Centros de Saúde (ACES) Baixo Mondego I. Esta avaliação pretende conciliar algumas das prioridades estabelecidas pelo Ministério da Saúde em 2011, em particular a secção sobre o "custo-efetividade, custo-benefício e custo-utilidade" que constam do mapa geral das necessidades de novos conhecimentos elaborado pela Coordenação Estratégica da Reforma dos CSP. Este trabalho envolve decisores reais participantes na avaliação de desempenho das unidades de CSP que monitorizam, recorrendo a uma nova abordagem. São apresentados neste texto apenas dois dos modelos analisados, que visam estabelecer uma ligação entre aspectos de avaliação como o acesso, os recursos utilizados, os serviços prestados e os resultados obtidos. A escolha dos fatores incluídos nos modelos foi resultado de discussões com o grupo de decisores (GD), constituído por profissionais do ACES Baixo Mondego I, com base no Plano de Desempenho (Coordenação Estratégica dos CSP [CECSP], 2010). A análise realizada utilizou dados relativos ao ano de 2010, compreendendo todas as 12 unidades prestadoras de cuidados de saúde que o ACES Baixo Mondego I monitoriza e acompanha.

A contribuição deste trabalho para o ACES Baixo Mondego I reside, essencialmente, na capacidade de usar múltiplos fatores num modelo integrado, dando a possibilidade de identificar as melhores práticas, as fontes de ineficiência, bem como as diferenças para as melhores práticas. Esta abordagem proporciona informações no sentido de serem introduzidas medidas corretivas e ajuda a estabelecer as metas de desempenho para o futuro.

O estudo da avaliação de eficiência na prestação de CSP tem sido alvo da atenção de vários autores, especialmente no Reino Unido. Na maioria dos trabalhos publicados para avaliar a prestação de serviços de saúde e qualidade de atendimento constata-se que são utilizados indicadores de *performance* (IPs) (Donabedian, 1980; Baker, 1992; Campbell et al., 1998; McColl et al., 1998; Van den Hombergh et al., 1998; Proctor e Campbell, 1999). Contudo, os modelos de avaliação de desempenho que são baseados num conjunto de IPs para comparar as entidades conduzem a análises pobres, uma vez que os diferentes IPs são analisados individualmente. Thanassoulis et al. (1996) e Giuffrida et al. (2000) apontaram as desvantagens do uso de IPs na avaliação de desempenho e propuseram técnicas alternativas, tais como a *Data Envelopment Analysis* (DEA). A DEA é uma abordagem não paramétrica, baseada em programação linear (PL), para avaliar a eficiência relativa de unidades organizacionais (*Decision Making Unit* - DMUs) semelhantes, gerando um indicador global de eficiência para cada unidade em avaliação, com base nos *inputs* que consome e nos *outputs* que produz. Desde meados de 1980 que esta metodologia é muitas vezes escolhida para avaliar a eficiência dos serviços de saúde (para revisões mais abrangentes ver Hollingworth, 2003; Amado & Dyson, 2008). Existem alguns estudos realizados sobre a eficiência na prestação de CSP que recorrem à técnica DEA (e.g., Huang & McLaughlin, 1989; Linna et al., 2003; Kirigia et al., 2004; Rosenman & Friesner, 2004; Rouse et al., 2011; Thanassoulis et al., 2012) e relativamente poucos estudos (Salinas-Jimenez & Smith, 1996; Wagner et al. 2003, Amado & Santos, 2009; Milliken et al., 2011; Nuti et al., 2011) incluem medidas de qualidade da assistência/atendimento na medição da eficiência.

Apesar de, nos últimos anos, se terem realizado muitos estudos de avaliação de eficiência em diferentes áreas utilizando DEA, poucos incorporam na análise as preferências dos gestores/decisores (ver Thanassoulis et al., 2004). A estrutura matemática dos modelos clássicos DEA permite total liberdade na escolha dos coeficientes de ponderação (pesos) dos *inputs* e dos *outputs*. Logo, a solução pode gerar coeficientes de ponderação nulos para variáveis importantes e eventualmente relações entre as variáveis, resultantes do próprio contexto do estudo, podem não ser consideradas. Como consequência, os resultados atingidos podem ser criticados pelo decisor, ficando comprometida a sua confiança na metodologia. Para ultrapassar este problema utilizam-se restrições aos coeficientes de ponderação com o objetivo de incorporar a opinião do decisor (Allen et al., 1997). As restrições mais comuns são as baseadas em restrições de razões entre pesos, na qual são definidos os limites inferiores e superiores atribuídos às razões entre os coeficientes de ponderação de duas quaisquer variáveis. Estas restrições definem uma região preferida da fronteira, onde são observadas as melhores práticas do modelo de produção em avaliação.

Este trabalho utiliza o método de duas fases desenvolvido por Gouveia et al. (2008), que é uma variante do modelo DEA aditivo com projeções orientadas (Ali et al., 1995), recorrendo a conceitos desenvolvidos no âmbito da MCDA com informação imprecisa (Athanassopoulos & Podinovski, 1997; Dias & Clímaco, 2000). Com as DMUs a assumir o papel de alternativas a serem comparadas, neste método os fatores de *input* e *output* são convertidos em funções de utilidade de acordo com informação de preferências fornecida pelos decisores e as funções de utilidade aditiva são utilizadas para agregar as utilidades (marginais) associadas a cada alternativa, com base na Teoria da Utilidade Multi-Atributo (*Multi-Attribute Utility Theory* - MAUT) (ver Keeney & Raiffa, 1976). Ao invés de deixar cada DMU escolher livremente os coeficientes de ponderação associados a estas funções, foram incluídas no modelo as preferências dos decisores no sentido de construir restrições aos coeficientes de ponderação seguindo alguns dos protocolos de diálogo recomendados (ver von Winterfeldt & Edwards, 1986). Neste método de duas fases, a medida de eficiência atribuída a cada DMU tem um significado intuitivo, correspondendo

a uma medida "*min-max regret*" (perda de utilidade). O modelo de Gouveia et al. (2008) foi neste estudo adaptado para incluir o conceito de super-eficiência (Gouveia et al., 2013), a fim de assinalar, entre as DMUs eficientes, quais o são com maior margem.

O método das duas fases de Gouveia et al. (2008), com as modificações para incluir o conceito de super-eficiência, é apresentado na secção 2. A secção 3 descreve brevemente o sistema de cuidados de saúde português. Na secção 4 são apresentados os dados, os modelos discutidos com o GD e os protocolos de diálogo utilizados para obter as suas preferências. Alguns resultados dos modelos propostos são exibidos e analisados na secção 5. A secção 6 destaca as propostas de melhoria para as unidades ineficientes e os novos desenvolvimentos do método apresentado na secção 2. As considerações finais são apresentadas na secção 7.

2. O método das duas fases baseado no modelo DEA aditivo com incorporação de preferências

As origens da DEA remontam aos trabalhos de Farrell (1957) sobre métodos para avaliação da produtividade, em que era reconhecida a necessidade de dispor de indicadores de eficiência global, e não de mera produtividade de fatores, que resultassem da combinação de múltiplos *inputs* e *outputs*. O primeiro modelo de DEA, assumindo rendimentos constantes à escala (*Constant Returns to Scale* - CRS), resultou do trabalho de Charnes, Cooper e Rhodes (Charnes et al., 1978), motivado pela avaliação de programas educacionais para estudantes desfavorecidos em escolas públicas dos Estados Unidos da América. A característica fundamental deste modelo (geralmente designado na literatura científica por modelo CCR ou *Engineering Ratio*) é, para cada entidade sob avaliação, a transformação da situação de múltiplos *outputs* / múltiplos *inputs* na situação de *output* virtual / *input* virtual. Para cada DMU, a razão entre o *output* virtual e o *input* virtual fornece uma medida de eficiência, a qual é função de um conjunto de multiplicadores que constituem as variáveis de um modelo de programação linear fraccionária (depois reduzido a um modelo de PL). As DMUs devem possuir um grau de controlo sobre os *inputs* consumidos e os *outputs* produzidos, ou seja,

autonomia de gestão, operando num ambiente relativamente homogêneo. Neste modelo espera-se uma variação proporcional dos produtos a partir da alteração de recursos em todos os níveis de escala.

A extensão do modelo CCR para o caso de rendimentos variáveis à escala (*Variable Returns to Scale - VRS*), foi proposta em Banker, Charnes e Cooper (Banker et al., 1984), modelo designado por BCC, onde basicamente se introduz uma restrição que garante que cada DMU em estudo será comparada com uma combinação convexa das restantes DMUs. De acordo com Ozcan (1993), pode dizer-se que o modelo VRS é a escolha mais adequada quando se pretende comparar unidades de saúde de tamanhos muito diferentes entre si, de modo que as unidades tenham como referência as outras com perfil semelhante.

Os diferentes modelos DEA procuram determinar quais das n DMUs formam a fronteira eficiente no sentido da eficiência de Pareto-Koopmans. Uma DMU será eficiente se, comparativamente às demais, tiver maior produção para quantidades fixas de recursos (orientação a *outputs*) e/ou utilizar menos recursos para gerar uma quantidade fixa de produtos (orientação a *inputs*). A DEA constrói uma fronteira empírica eficiente, ou "superfície envolvente" de máxima eficiência (Ali & Seiford, 1993). Isto permite a identificação de unidades de referência (*benchmarks*) cujos valores de eficiência servem como modelo para as restantes unidades envolvidas pela fronteira de eficiência.

Charnes et al. (1985) propuseram o modelo DEA aditivo como uma alternativa para o modelo BCC, que também considera que os rendimentos podem ser variáveis à escala, mas não requer uma escolha entre modelo orientado a *inputs* e modelo orientado a *outputs*.

A distinção entre modelo DEA aditivo e os modelos orientados reside no facto destes últimos terem um mecanismo de projecção em duas etapas. Nos modelos orientados, numa primeira etapa todos os fatores são reduzidos ou aumentados na mesma proporção até à superfície envolvente, e numa segunda etapa são calculadas as folgas que dão o ponto projetado na fronteira eficiente. Por sua vez, o modelo DEA aditivo utiliza somente a segunda etapa. O valor obtido na primeira etapa de modelos orientados fornece uma medida da

ineficiência, entre 0 e 1, para as DMUs ineficientes de acordo com a orientação utilizada.

Seja $\{DMU_j: (j = 1, \dots, n)\}$ o conjunto das n DMUs a ser avaliadas. Cada DMU j consome a quantidade $x_{ij} (i = 1, \dots, m)$ do *input* i e produz a quantidade $y_{rj} (r = 1, \dots, p)$ do *output* r . O modelo aditivo a resolver para cada DMU tem a formulação primal (1):

$$\begin{aligned}
 \min_{\lambda, s^-, s^+} z_k &= -\left(\sum_{i=1}^m s_i^- + \sum_{r=1}^p s_r^+\right) \\
 \text{s. a: } -\sum_{j=1}^n \lambda_j x_{ij} - s_i^- &= -x_{ik}, \quad i = 1, \dots, m \\
 \sum_{j=1}^n \lambda_j y_{rj} - s_r^+ &= y_{rk}, \quad r = 1, \dots, p \\
 \sum_{j=1}^n \lambda_j &= 1 \\
 \lambda_j, s_i^-, s_r^+ &\geq 0, \quad j = 1, \dots, n; i = 1, \dots, m; r = 1, \dots, p
 \end{aligned} \tag{1}$$

O modelo DEA aditivo não devolve uma medida para as DMUs ineficientes. Este modelo apenas mede o excesso de *inputs* (s^-) e o déficit de *outputs* (s^+), em que uma DMU k está a operar quando comparada com as DMUs que operam na fronteira eficiente. Se o valor obtido for negativo, então a DMU em análise está a operar de forma ineficiente em alguns fatores. Este valor é simétrico da soma das distâncias em cada dimensão à superfície envolvente (distância L_1).

O método de duas fases desenvolvido por Gouveia et al. (2008) é uma variante do modelo DEA aditivo com projeções orientadas (Ali et al., 1995), que incorpora as preferências do decisor utilizando conceitos da MAUT com informação imprecisa (Dias & Clímaco, 2000).

O objetivo da MAUT é encontrar uma única expressão que represente a atratividade global de cada alternativa, de acordo com as preferências do decisor, tendo em consideração os múltiplos critérios. É também propósito desta teoria simplificar a construção de funções de utilidade multicritério deixando que o decisor foque a sua atenção num critério de cada vez, e posteriormente na agregação dos vários critérios, em vez de avaliar diretamente a utilidade global.

A eliciação de preferências é realizada interativamente, em pequenos passos, de forma construtiva. von Winterfeldt e Edwards (1986) apresentam detalhadamente várias técnicas de questionar o decisor no sentido de construir as funções de utilidade marginais (parciais) consistentes com as respostas obtidas.

Seja $u_c(a_j)$ a utilidade de uma alternativa a_j de acordo com o critério c ($c = 1, \dots, q$). A função de utilidade aditiva agrega estas funções através de uma soma ponderada na qual são usados q coeficientes de escala, w_1, \dots, w_q : $U(a_j) = \sum_{c=1}^q w_c u_c(a_j)$ (ver Keeney & Raiffa, 1976). A quantidade $u_c(a_j)$ pode ser interpretada como uma medida de desempenho da alternativa a_j no critério c . Estas funções são definidas para cada critério convertendo a escala de desempenhos originais na escala de utilidades $[0,1]$. Por sua vez, w_c é o coeficiente de ponderação que reflete a importância da função u_c ($c = 1, \dots, q$). Os coeficientes de ponderação são normalizados de forma a que $\sum_{c=1}^q w_c = 1$, com $w_c \geq 0, \forall c = 1, \dots, q$, podendo ainda ser acrescentadas pelo decisor outras restrições, caso se pretenda impedir combinações de valores para estes coeficientes que não estejam alinhados com os objetivos da organização.

Este procedimento ultrapassa o problema das escalas associado ao modelo DEA aditivo, uma vez que todas as medidas de *input* e *output* são convertidas em utilidade. Além disso, os coeficientes de ponderação usados na agregação adquirem um significado específico: são os coeficientes de escala das funções de utilidade. Os coeficientes de ponderação são valores escolhidos por cada DMU, a fim de se tornar a melhor DMU (se possível) ou então no sentido de minimizar a diferença de utilidade para a melhor DMU, ao invés de serem fixados previamente como no modelo de Ali et al. (1995).

A teoria de decisão clássica oferece algumas regras (tais como *max-min* e *min-max regret*) que podem ser aplicadas a situações onde as funções de utilidade são incertas.

No desenvolvimento do método das duas fases utilizou-se a regra *min-max regret*, referente ao conceito de perda de utilidade de escolher uma alternativa face a outra. A ideia é determinar para cada alternativa o potencial “arrependimento” associado à sua escolha, observando quão melhor poderia outra alternativa ser. Para um dado vetor de coeficientes de ponderação, a

perda de utilidade de escolher a alternativa a_j em vez de outra alternativa $a_j \neq a_k$ é dada por (2):

$$R(a_k, a_j) = \max_{j \neq k} U(a_j) - U(a_k) \quad (2)$$

A perda de utilidade máxima associada a uma dada alternativa a_k , $R_{max}(a_k)$, é dada pela máxima perda de oportunidade associada à escolha dessa alternativa (3):

$$R_{max}(a_k) = \max_w \{(\max_{j \neq k} U(a_j)) - U(a_k)\} \quad (3)$$

A decisão ótima *min-max regret* minimiza a perda no pior caso de acordo com possíveis realizações do vetor de coeficientes w , i.e., alternativas com $R_{max}(a_k)$ mínimo são preferidas. Assim, a medida de eficiência atribuída a cada DMU ganha um significado intuitivo: corresponde a uma medida "*min-max regret*" (perda de utilidade).

Considerando que as alternativas a avaliar de acordo com q ($q = m+p$) critérios são as DMUs, definem-se q funções de utilidade de modo a que o pior nível da escala tenha o valor 0 e o melhor nível da escala tenha o valor 1. Por outras palavras, quanto melhor (neste caso menor) for o desempenho associado a um fator de *input* maior será a sua utilidade (mais próxima estará de 1); para o caso dos *outputs*, quanto maior for o desempenho maior será a sua utilidade. Assim, depois de todos os fatores terem sido convertidos em utilidades, passam a ser considerados como *outputs* a maximizar. Para cada alternativa (DMU), atendendo ao modelo aditivo MAUT, a utilidade obtida é dada por: $U(DMU_j) = \sum_{c=1}^q w_c u_c(DMU_j)$, onde $w_c \geq 0, \forall c = 1, \dots, q$ e $\sum_{c=1}^q w_c = 1$. Os coeficientes de escala w_1, \dots, w_q são os coeficientes de ponderação das funções de utilidade e traduzem as taxas de compensação (*trade-offs*) da utilidade do decisor entre os vários fatores.

Para atender ao conceito de super-eficiência, introduzido em DEA por Andersen e Petersen (1993), o método de duas fases de Gouveia et al. (2008) foi modificado no sentido de dar a possibilidade de discriminar as unidades eficientes (Gouveia et al., 2013). Assim, o problema (4) é resolvido (Fase 1):

$$\begin{aligned}
& \min_{d_k, w} d_k \\
& \text{s. a: } \sum_{c=1}^q w_c u_c(DMU_j) - \sum_{c=1}^q w_c u_c(DMU_k) \leq d_k, \quad j = 1, \dots, n; j \neq k \\
& \quad \sum_{c=1}^q w_c = 1 \\
& \quad w_c \geq 0, c = 1, \dots, q
\end{aligned} \tag{4}$$

O valor ótimo d_k^* indica a distância definida pela diferença de utilidade para a melhor de todas as DMUs (excluindo a que está em avaliação).

O propósito desta abordagem é, para a DMU k , calcular o vetor w de ponderação que minimiza a distância (a diferença de utilidade) da unidade em causa para a melhor (note-se que a melhor alternativa dependerá também de w), excluindo-se ela própria do conjunto de referência. Depois, no caso de ser uma DMU ineficiente ou fracamente eficiente, encontra um alvo (par) eficiente (Fase 2). Os detalhes deste processo são os seguintes:

Fase 1: Converter os fatores de *input* e *output* em escalas de utilidade. Calcular a medida de eficiência, d_k^* , de cada DMU, $k = 1, \dots, n$, e o vetor de ponderação correspondente.

Fase 2: Se $d_k^* \geq 0$, então resolver o problema "aditivo ponderado" (5), usando o vetor de ponderação ótimo resultante da Fase 1, w_k^* , e determinar o correspondente ponto projetado da DMU em análise.

$$\begin{aligned}
& \min_{\lambda, s} z_k = - \sum_{c=1}^q w_c^* s_c \\
& \text{s. a: } \sum_{j=1, j \neq k}^n \lambda_j u_c(DMU_j) - s_c = u_c(DMU_k), \quad c = 1, \dots, q \\
& \quad \sum_{j=1, j \neq k}^n \lambda_j = 1 \\
& \quad \lambda_j \geq 0, s_c \geq 0, \quad j = 1, \dots, k-1, k+1, \dots, n; c = 1, \dots, q
\end{aligned} \tag{5}$$

Se o valor ótimo d_k^* da função objetivo em (4) é não positivo, então a DMU k em avaliação é eficiente (se o valor ótimo for nulo a DMU poderá ser fracamente eficiente). Caso contrário, é ineficiente e d_k^* é a diferença mínima de utilidade para a melhor DMU (ou seja, a DMU com maior utilidade global). Contudo, com a modificação feita no método original, é possível discriminar as unidades eficientes. As variáveis $\lambda_1, \dots, \lambda_j, j = 1, \dots, k-1, k+1, \dots, n$, definem uma combinação convexa das n DMUs. As DMUs eficientes (podendo ser apenas uma) que definem a combinação convexa com $\lambda_j > 0$ são chamadas de "pares" da DMU k em avaliação. A combinação convexa corresponde a um

ponto na fronteira eficiente, melhor do que a DMU k com uma diferença de valor de s_c (folga) em cada critério c .

O modelo (5) admite alvos ótimos alternativos, cada um dos quais correspondendo a uma forma diferente de anular o valor da diferença d_k^* . Esses alvos correspondem a diferentes projeções sobre a fronteira eficiente. Existe a possibilidade de escolher quais os fatores que se permite alterar e limitar a melhoria máxima que se pode exigir em cada fator para atingir o valor global de utilidade do alvo, através da inserção de restrições às folgas s_c . O objetivo é evitar a fixação de alvos irrealistas, permitindo que as propostas de melhoria das unidades ineficientes sejam concordantes com os requisitos do decisor.

3. O sistema de saúde português – Cuidados de saúde primários

Os CSP são a primeira e principal forma de contacto da população com o sistema de saúde. Em 1979, o SNS foi criado com o objectivo de defender o direito à saúde e o dever de defesa e protecção da saúde da população. O SNS forneceria um serviço nacional, universal, integral e gratuito, assegurado pelo Estado. A especialidade de medicina geral e familiar foi também criada e implementada. Apesar da maioria da população ser utilizadora do SNS, há vários subsistemas de seguros de saúde, mas a população coberta por estes também tem acesso aos serviços do SNS, traduzindo-se, muitas vezes, numa dupla cobertura.

Sucessivas tentativas de reforma do SNS conduziram ao entendimento de que não foi dada a devida atenção ao acesso aos cuidados de saúde, sendo necessário projetar políticas de saúde que se traduzam em soluções técnicas que realmente funcionem. É igualmente importante que a base política de qualquer reforma seja alicerçada numa análise rigorosa das questões económicas, sociais, culturais e políticas (Observatório Português dos Sistemas de Saúde [OPSS], 2008).

Em 2005 começou uma nova fase no sector da saúde, especialmente em relação aos CSP, focando em particular a reorganização dos centros de saúde, com a criação de Unidades de Saúde Familiar (USF), constituídas por equipas multidisciplinares com autonomia técnica e funcional, e de outras soluções

mais adequadas para a prestação de cuidados de saúde primários. O impacto desta reforma está a ser acompanhado, o que justifica a necessidade de avaliar os seus resultados, em particular no que diz respeito às melhores práticas em CSP.

Do ponto de vista organizacional dos CSP, pode referir-se a existência, em 2005, de 351 centros de saúde, 1823 extensões de saúde e 254 Serviços de Atendimento Permanente (SAPs) em Portugal Continental, de acordo com o Instituto de Gestão Informática e Financeira da Saúde (2006). Em termos de modelo de contratualização, constatava-se a ausência de um papel claro e bem definido de um agente que desse resposta às necessidades dos diversos atores com responsabilidades no sector da saúde, especialmente no que diz respeito à negociação de objetivos e metas. De acordo com um estudo da Organização para a Cooperação e Desenvolvimento Económico (2006), os gastos com a saúde aumentariam no futuro, estimando-se representar 13.1% do PIB em 2050, se não fossem implementadas reformas para contrariar esta tendência.

A prestação de serviços de saúde é caracterizada por um grande número de utilizadores sem acesso a médico de família, especialmente na área geográfica do litoral e nas grandes áreas urbanas. Os cidadãos podem escolher livremente o centro de saúde onde se inscrevem, bem como o seu profissional de medicina geral e familiar. Este médico fica designado como médico de família do utente para todas as consultas. No entanto, em 2005, cerca de 10.6% dos utentes do SNS não tinham médico de família (OPSS, 2007).

Durante todo o ano de 2007 houve um compromisso do Ministério da Saúde para a criação de mais USF e, a 31 de dezembro de 2007, já se encontravam 104 USF em funcionamento.¹

Foram estabelecidos modelos de acompanhamento, avaliação e contratualização, com padrões de exigência elevados. Estes conduziram à criação de um clima de confiança entre os profissionais de saúde envolvidos e as Administrações Regionais de Saúde (ARS). O ambiente criado mobilizou cerca de 1300 médicos, 1300 enfermeiros e 950 funcionários administrativos, que aceitaram o desafio e se disponibilizaram para trabalhar em USF. Para além destas mudanças organizacionais, a lista de utentes inscritos, em média,

¹ Dados retirados do sítio www.mcsp.min-saude.pt, consultado em 15 de janeiro de 2012.

aumentou para cerca de 1750 utentes por profissional de medicina geral e familiar (OPSS, 2008).

A 22 de fevereiro de 2008 foi publicado o Decreto-Lei n.º 28/2008, que regulamenta a criação dos ACES. Estes são definidos como “*serviços públicos de saúde com autonomia administrativa, constituídos por várias unidades funcionais, que agrupam um ou mais centros de saúde, e que têm por missão garantir a prestação de cuidados de saúde primários à população de determinada área geográfica*”. A constituição dos ACES pressupõe a sua organização em unidades funcionais, nomeadamente em USF, unidades de cuidados de saúde personalizados (UCSP - integram os restantes médicos de família e outros profissionais, mas não estão organizados em USF), unidades de cuidados na comunidade, unidades de saúde pública e unidades de recursos assistenciais partilhados. O ACES também prevê a existência de unidades de apoio à sua governação, como é o caso do Conselho da Comunidade, o Conselho Clínico, a Unidade de Apoio à Gestão e o Gabinete do Cidadão. Importa realçar que a autonomia do ACES é administrativa e está intimamente relacionada com a existência de um contrato-programa, a ser estabelecido entre o ACES e a ARS, Instituto Público.

De acordo com o Decreto-Lei n.º 222/2007, publicado em 29 de maio, as cinco ARS existentes “têm por missão garantir à população da respectiva área geográfica de intervenção o acesso à prestação de cuidados de saúde de qualidade, adequando os recursos disponíveis às necessidades em saúde e cumprir e fazer cumprir o Plano Nacional de Saúde na sua área de intervenção”.

Em 2010, pela primeira vez, os Diretores Executivos/Conselhos Clínicos, com o apoio dos Departamentos de Contratualização das ARS, desencadearam a contratualização interna com as USF, já iniciada anteriormente, e o mesmo modelo foi adaptado para as UCSP.

Conforme consta no contrato-programa de 2010, para um futuro a curto prazo “*espera-se que o processo de contratualização e a sofisticação da alocação dos recursos permita uma efetiva autonomização e responsabilização dos ACES pela efetividade dos cuidados de saúde prestados à população*”.

4. Dados e modelos desenvolvidos atendendo às preferências do Grupo de Decisores

Esta secção descreve a avaliação de desempenho das 12 unidades prestadoras de cuidados de saúde monitorizadas pelo ACES Baixo Mondego I, com dados relativos a 2010. As unidades a avaliar são USF e UCSP, uma vez que todas elas estão no mesmo contrato-programa estabelecido entre o ACES e a ARS e prosseguem os mesmos objetivos. A avaliação de desempenho destas unidades é realizada periodicamente pelo ACES, tendo como base os indicadores que constam no Plano de Desempenho, sem atender a diferenças existentes a nível organizacional. Portanto, e a pedido do GD, manteve-se o mesmo critério na avaliação proposta neste estudo, podendo-se considerar que as unidades são homogéneas.

A monitorização e acompanhamento destas unidades assenta num conjunto de 15 indicadores agrupados em quatro grandes áreas: acesso, desempenho assistencial, qualidade percecionada e desempenho económico, cujas metas são negociadas entre a unidade e o Diretores Executivos/Conselhos Clínicos, tendo em consideração os objetivos de saúde estabelecidos pelo ACES e pela ARS. Um bom desempenho por parte das unidades poderá traduzir-se na atribuição de recompensas de vária ordem.

Os dados analisados foram extraídos de diversos sistemas de informação, desde os que são utilizados para monitorizar as atividades assistenciais até aos do departamento de contabilidade e gestão. A fim de incluir na análise um fator que reflete a qualidade percecionada de serviço, foi utilizada informação sobre a satisfação dos utentes com a qualidade do serviço, proveniente de um inquérito (*Avaliação sobre o funcionamento das Unidades de Saúde do ACES*, 2010). Este abrange 19 perguntas sobre acessibilidade aos serviços, informação e atendimento, instalações das unidades de saúde e dados sociodemográficos. O inquérito não considerou um indicador de satisfação global, mas foi entendido pelo GD que a pergunta "*quando adoecer de repente, consegue consulta no médico de família no próprio dia?*" (resposta sim/não) seria uma boa aproximação da satisfação dos utentes com a qualidade do serviço.

Face ao objetivo de avaliação do desempenho das 12 DMUs (USF e UCSP) monitorizadas pelo ACES Baixo Mondego I, foi explicada ao GD a metodologia

proposta por Gouveia et al. (2008). O GD aceitou esta abordagem que considerou muito útil, dada a possibilidade de usar múltiplos fatores num modelo integrado e a capacidade de identificar as melhores práticas, as fontes de ineficiência, as diferenças relativamente às melhores práticas, bem como as oportunidades de melhoria.

4.1. Definição dos modelos

Depois de traçados os objetivos do estudo, foram definidos os fatores de *input* e *output* para cada modelo, atendendo aos indicadores das principais áreas do Plano de Desempenho, que eram essencialmente rácios. Como resultado de reuniões surgiram vários modelos diferentes, pois vários entendimentos possíveis decorriam das diferentes perspectivas de avaliação utilizadas. No entanto, neste trabalho apenas serão apresentados dois desses modelos.

Com o objetivo de combinar desempenho assistencial, qualidade percebida e desempenho económico, foi desenvolvido um modelo de eficiência face a custos. Neste, os custos associados à prestação de cuidados de saúde primários foram considerados como *inputs*; como *outputs* foram incluídos os fatores considerados como a principal atividade dos centros de saúde, juntamente com o fator relativo à satisfação dos utentes com a qualidade do serviço de saúde prestado. Os *outputs* (à exceção do último referido) são separados por tipo de consulta e por grupo etário dos utentes, a fim de controlar o “*case-mix*” (Tabela 1). Este modelo foi chamado de “eficiência dos serviços de saúde prestados atendendo às características do utente”, a partir de agora referido como modelo 1.

Uma vez que no modelo 1 não está refletido o trabalho de enfermagem, houve a necessidade por parte do GD de criar um segundo modelo de eficiência (modelo 2). Este modelo, considerado mais abrangente, alia os vários eixos de avaliação mencionados no Plano de Desempenho, expressando em simultâneo todo o trabalho médico, de enfermagem e administrativo. Concordou-se com a inclusão de nove fatores para a construção deste modelo (ver Tabela 2), cujos *inputs* são os indicadores económicos do Plano de Desempenho, os primeiros quatro *outputs* estão relacionados com o acesso

(número de utentes inscritos com pelo menos uma consulta no ano de referência) e desempenho assistencial e o último *output* é relativo à qualidade percebida. Este modelo fornece ao GD uma perspectiva diferente das unidades em avaliação e foi considerado muito importante, uma vez que as recompensas institucionais são dependentes do desempenho das unidades em alguns destes fatores, nomeadamente o número de domicílios médicos e de enfermagem.

Tabela 1. Fatores do modelo 1.

<i>Inputs</i>	<i>Outputs</i>
X_{MCDT} : custo total com meios complementares de diagnóstico e tratamento faturados ao SNS	y_{CPF} : número de consultas médicas de planeamento familiar
X_{MED} : custo total com medicamentos faturados ao SNS	y_{CSM} : número de consultas de saúde materna
X_{RH} : custo total com recursos humanos	y_{CSI} : número de consultas de saúde infantil (idade 0-18)
X_{OC} : custos totais com medicamentos, materiais de consumo clínico e outros	y_{CSA} : número de consultas de saúde a adultos (idade 19-64)
	y_{CSS} : número de consultas de saúde a seniores (idade superior ou igual a 65)
	y_{SQ} : satisfação com a qualidade dos serviços

Amplamente discutidos com o GD e em concordância com os indicadores estabelecidos no Plano de Desempenho, estes *inputs* e *outputs* são típicos deste tipo de aplicação DEA (e.g. Amado & Santos, 2009).

4.2. Construção das funções de utilidade dos fatores

Existem várias técnicas para obter informações de preferência de um Grupo de Decisores/Decisor com o propósito de construir funções de utilidade compatíveis com as suas respostas (ver von Winterfeldt & Edwards, 1986; Goodwin & Wright 1998); no entanto, as perguntas para esse fim devem ser

enquadradas no contexto de cada estudo. No método de Gouveia et al. (2008) o objetivo de converter os fatores numa escala de utilidade (funções de utilidade não-lineares) é o de espelhar as preferências do GD. Para construir as funções de utilidade seguiu-se um protocolo de diálogo já utilizado por Almeida e Dias (2012).

Tabela 2. Fatores do modelo 2.

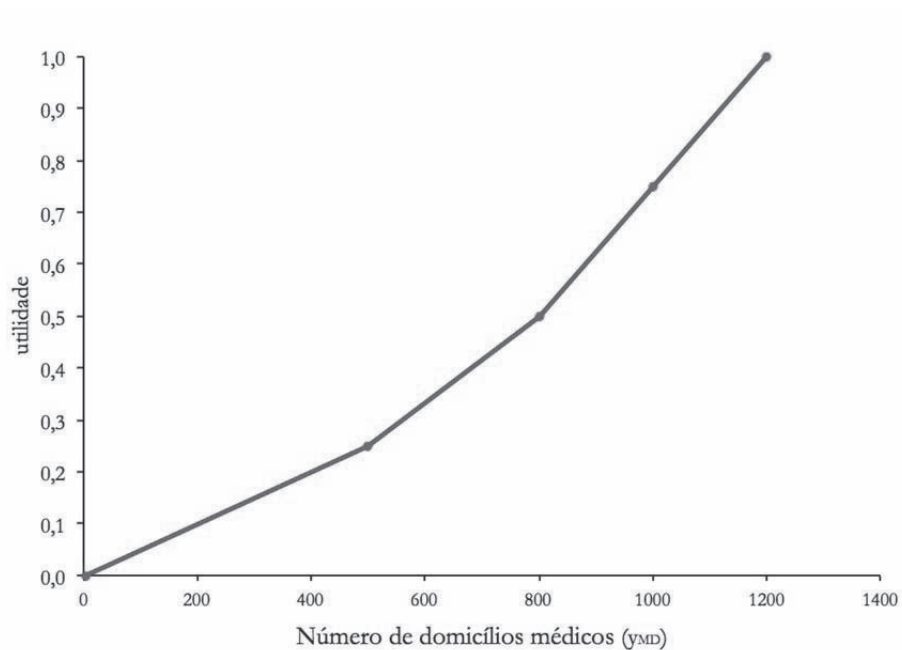
<i>Inputs</i>	<i>Outputs</i>
x_{MCDT} : custo total com meios complementares de diagnóstico e tratamento faturados ao SNS	y_{DM} : número de domicílios médicos y_{CM} : número de consultas médicas na unidade
x_{MED} : custo total com medicamentos faturados ao SNS	y_{DE} : número de domicílios de enfermagem
x_{RH} : custo total com recursos humanos	y_{CE} : número de consultas de enfermagem na unidade
x_{OC} : custos totais com medicamentos, materiais de consumo clínico e outros	y_{SQ} : satisfação com a qualidade dos serviços

O protocolo de diálogo baseou-se na comparação entre o mérito de aumentar um *output* (ou diminuir um *input*) de a para b contra o aumento do mesmo *output* (ou diminuição do mesmo *input*) de a' para b' , mantendo todos os outros níveis de desempenho iguais, e pedindo ao GD para ajustar um destes quatro valores de tal forma que o aumento de mérito seja aproximadamente igual. Para fazer esta conversão será apenas assumido que as funções são contínuas.

Por exemplo, considerando o *output* y_{DM} a questão colocada ao GD foi: “Tem mais mérito aumentar o número de domicílios médicos de 3 para 600 ou de 600 para 1200, mantendo todos os outros desempenhos iguais?” A resposta foi: “tem mais mérito aumentar de 600 para 1200”. Então o nível foi ajustado e a questão foi reformulada como: “Tem mais mérito aumentar o número de domicílios médicos de 3 para 800 ou de 800 para 1200, mantendo todos os outros desempenhos iguais?” A resposta foi: “tem igual mérito”. Um procedimento análogo foi utilizado para dicotomizar os intervalos de mérito [3,800] e [800,1200].

As funções de utilidade, para a maioria dos fatores, foram representadas por aproximação a funções lineares definidas por ramos. Quando as respostas do GD se conseguiram ajustar a curvas pré-definidas (como funções logarítmicas) as funções utilidade foram representadas por essas funções conhecidas.

Os intervalos encontrados foram escolhidos com o intuito de incluir as gamas de desempenho observadas para cada fator, mais (ou menos) uma margem de tolerância máxima (neste caso de 20%).



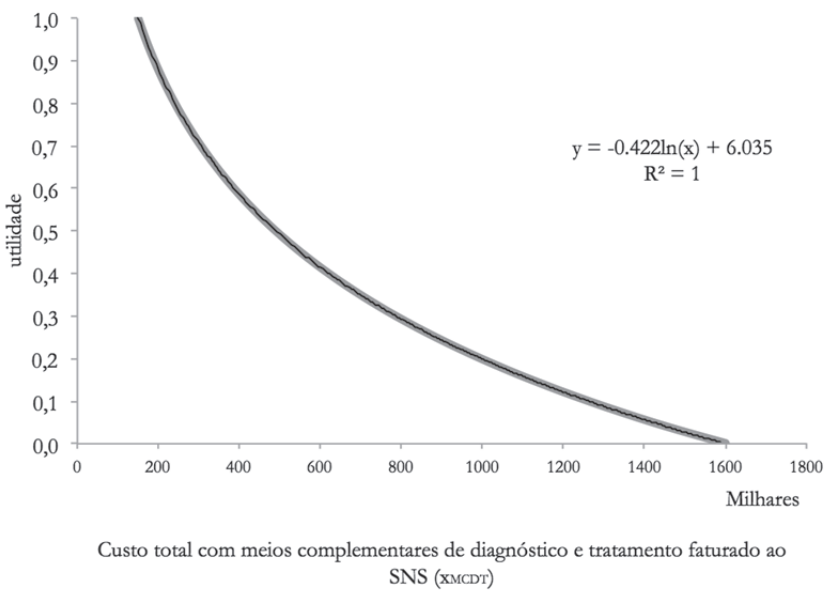
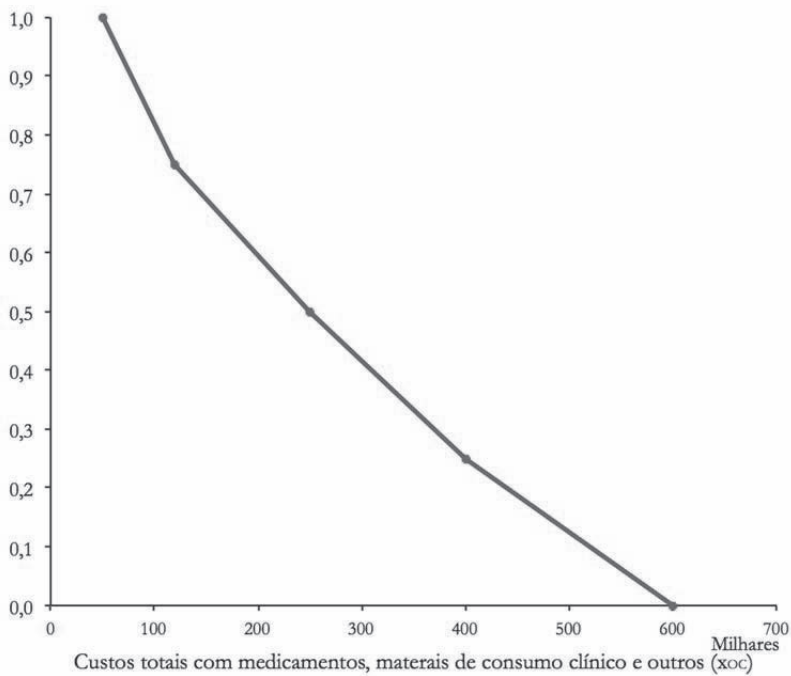


Figura 1. Três das funções de utilidade eliciadas.

As funções de utilidade lineares definidas por ramos para alguns dos fatores estão exibidas na Figura 1. Por exemplo, para o número de domicílios médicos:

$u_{DM}(1200) - u_{DM}(1000) = u_{DM}(1000) - u_{DM}(800) = u_{DM}(800) - u_{DM}(500) = u_{DM}(500) - u_{DM}(3)$, mantendo-se todos os outros níveis de desempenho iguais. As outras funções de utilidade para os restantes fatores *output* dos modelos (y_{CPF} , y_{CSM} , y_{CSI} , y_{CSA} , y_{CSS} , y_{CM} , y_{DE} , y_{CE} , e y_{SQ}) são semelhantes à função de utilidade y_{DM} .

Para os custos x_{MCDT} , x_{MED} , e x_{RH} , as funções de utilidade foram obtidas realizando o ajustamento de uma função conhecida às preferências reveladas pelo GD.

4.3. Restrições aos coeficientes de ponderação

Para os modelos apresentados houve a necessidade de incorporar restrições aos coeficientes de ponderação porque, caso contrário, não era possível alcançar uma classificação das unidades com discriminação aceitável (veja-se Dyson et al., 2001). Isto foi explicado ao GD, acrescentando que a introdução de restrições aos coeficientes de ponderação nos modelos traduz melhor os objetivos da organização e, portanto, garante resultados significativamente mais próximos daquilo que o GD considera como sendo as melhores práticas.

Existem na literatura várias abordagens para a definição de restrições aos coeficientes de ponderação, podendo a especificação destes de forma adequada revelar-se um desafio (Salo & Hämäläinen, 2001; Podinovski, 2004).

A técnica de *swings* começa pela construção de duas hipóteses extremas, B e G; a primeira diz respeito ao pior desempenho (utilidade 0) em todos os critérios e a segunda corresponde ao melhor desempenho (utilidade 1). O protocolo de diálogo consiste em consultar o decisor no sentido deste olhar para potenciais ganhos ao passar de B para G em cada critério e, em seguida, decidir quais os critérios que prefere mudar para a hipótese G. Supõe-se que a transição da hipótese B para a hipótese G, num determinado critério, vale 100 unidades numa escala hipotética. Em seguida, é solicitada ao decisor a atribuição de um valor (<100) ao segundo critério movido para G, depois ao

terceiro critério e assim sucessivamente, até que o último critério é movido para G.

O procedimento utilizado neste trabalho foi: primeiro obter uma ordenação de coeficientes de ponderação (baseado na técnica de *swings*) e, depois, estabelecer um limite para a razão entre os coeficientes de ponderação que ficaram em primeiro e em último lugar (pergunta de *trade-off*), por forma a evitar pesos nulos.

Ao incluir os coeficientes de ponderação, a Fase 1 e a Fase 2 (ver secção 2) passam a assumir outras formulações: às restrições do programa linear da Fase 1 acrescentam-se as restrições aos pesos, ao passo que no programa linear da Fase 2 é necessário permitir às folgas s_c ter valores negativos, retirando-se a restrição $s_c \geq 0$ (ver detalhes em Almeida & Dias, 2012).

5. Resultados

A implementação computacional do método, bem como das suas extensões, foram realizadas recorrendo a uma folha de cálculo.

O analista em conjunto com o GD construíram dois modelos (1 e 2) para avaliar a eficiência das unidades de CSP. A Tabela 3 apresenta os dados originais e as utilidades correspondentes para o modelo 1.

Tabela 3. Desempenhos das DMUs nas escalas originais e na escala de utilidades para o modelo 1.

Fatores nas escalas originais										
DMU	<i>X_{MCDT}</i>	<i>X_{MED}</i>	<i>X_{RH}</i>	<i>X_{OC}</i>	<i>Y_{CPF}</i>	<i>Y_{CSM}</i>	<i>Y_{CSI}</i>	<i>Y_{CSA}</i>	<i>Y_{CSS}</i>	<i>Y_{SQ}</i>
1	362900	1114935	1398033	123382	845	323	5133	25380	17149	92
2	463549	1192330	1050558	128606	449	311	3197	19983	14783	74
3	466561	1113734	1589818	101750	85	76	691	5136	2434	81
4	277140	1321779	1001297	152974	1628	605	4934	19283	13811	83
5	232565	1147295	1101200	141864	732	363	1668	17775	12108	62
6	613535	2491923	1724028	218890	2534	782	5068	29422	18854	75
7	1332317	4365605	2368401	360011	3726	1081	8073	54200	40307	77
8	701121	2289334	2225321	491140	2641	1142	4530	41247	27858	80
9	314090	1196353	1054564	65416	666	502	4513	21909	13476	91
10	606148	2954054	1348651	203098	3353	662	5068	29422	18854	54
11	1007909	3350321	1933791	371468	4763	1084	6587	42838	30037	78
12	708936	2793686	2227392	278936	3087	1052	5328	35475	25593	74
Fatores em escalas de utilidade										
DMU	<i>U_{MCDT}</i>	<i>U_{MED}</i>	<i>U_{RH}</i>	<i>U_{OC}</i>	<i>U_{CPF}</i>	<i>U_{CSM}</i>	<i>U_{CSI}</i>	<i>U_{CSA}</i>	<i>U_{CSS}</i>	<i>U_{SQ}</i>
1	0.633	0.824	0.572	0.743	0.084	0.123	0.392	0.256	0.210	0.750
2	0.529	0.789	0.788	0.733	0.042	0.117	0.193	0.190	0.178	0.400
3	0.527	0.824	0.475	0.815	0.003	0.010	0.014	0.014	0.006	0.521
4	0.746	0.735	0.825	0.687	0.168	0.253	0.367	0.182	0.164	0.563
5	0.820	0.809	0.753	0.708	0.072	0.141	0.083	0.164	0.140	0.220
6	0.411	0.406	0.413	0.560	0.271	0.364	0.384	0.324	0.234	0.417
7	0.084	0.115	0.173	0.317	0.457	0.568	0.759	0.737	0.633	0.450
8	0.355	0.450	0.220	0.136	0.288	0.618	0.316	0.521	0.381	0.500
9	0.694	0.787	0.785	0.945	0.065	0.205	0.314	0.213	0.159	0.729
10	0.416	0.318	0.599	0.590	0.399	0.289	0.384	0.324	0.234	0.140
11	0.202	0.253	0.326	0.298	0.659	0.570	0.573	0.547	0.417	0.467
12	0.350	0.347	0.219	0.452	0.357	0.543	0.416	0.425	0.343	0.400

Como anteriormente foi referido, para estes dois modelos, dada a relação entre o número de fatores e o número de unidades em avaliação, foi necessário incorporar restrições aos coeficientes de ponderação.

De acordo com o procedimento explicado na sub-secção 4.3, as restrições aos coeficientes de ponderação foram obtidas pedindo ao GD para comparar os "swings" da utilidade de 0 para 1 para o modelo 1, como descrito na Tabela 4.

O GD foi convidado a considerar uma unidade com o desempenho no nível 0 para todos os fatores e perguntou-se: "se pudesse melhorar um e apenas um fator para o nível 1, qual seria?". A resposta do GD foi unânime: "custo total com medicamentos faturados ao SNS". Isto permite a ilação de que w_{MED} é o maior coeficiente de escala. Ao repetir esta pergunta sucessivamente para os demais fatores, e depois de alguma discussão, o GD confluía na seguinte ordenação para os coeficientes de ponderação:

$$w_{MED} \geq w_{MCDT} \geq w_{RH} \geq w_{SQ} \geq w_{CPF} \geq w_{CSI} \geq w_{CSM} \geq w_{OC} \geq w_{CSA} \geq w_{CSS}.$$

Tabela 4. Desempenhos extremos associados com os níveis de utilidade 0 e 1 (modelo 1).

Nível de utilidade	X_{MCDT}	X_{MED}	X_{RH}	X_{OC}	Y_{CPF}	Y_{CSM}	Y_{CSI}	Y_{CSA}	Y_{CSS}	Y_{SQ}
$u(.)=0$	1600000	5500000	3000000	600000	60	55	500	4000	2000	40
$u(.)=1$	150000	800000	800000	50000	6000	1500	10000	65000	50000	100

Com o objetivo de evitar valores nulos nos coeficientes de ponderação, foi feita ao GD uma pergunta de *trade-off* para limitar a relação dos coeficientes de ponderação classificados na primeira (w_{MED}) e última posição (w_{CSS}): "Qual seria o maior valor z em euros correspondente ao custo total com medicamentos faturados ao SNS que permitiria a uma unidade com 50000 consultas de saúde a seniores realizadas e custo total com medicamentos faturados ao SNS associado de 5500000 euros ser considerada como tendo mais mérito do que uma unidade com 2000 consultas de saúde a seniores realizadas e com um custo total com medicamentos faturados ao SNS de z euros?". A resposta permite construir a seguinte desigualdade:

$$w_{CSS} u(50000) + w_{MED} u(5500000) \geq w_{CSS} u(2000) + w_{MED} u(z).$$

O limite obtido foi $z = 2500000$. Com o aumento deste valor o GD começava a hesitar se a desigualdade se manteria. Substituindo esse valor de z na expressão anterior: $w_{MED} \leq 2.47 w_{CSS}$.

Os resultados para o modelo 2 estão representados nas Tabela 5 e 6.

Tabela 5. Resultados da Fase 1, com restrições aos coeficientes de ponderação (modelo 1).

Fase 1											
DMU	d^*	w^*_{MCDT}	w^*_{MED}	w^*_{RH}	w^*_{OC}	w^*_{CPF}	w^*_{CSM}	w^*_{CSI}	w^*_{CSA}	w^*_{CSS}	w^*_{SQ}
1	0.0223	0.087	0.215	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087
2	0.0813	0.171	0.171	0.171	0.069	0.069	0.069	0.069	0.069	0.069	0.069
3	0.1426	0.087	0.215	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087
4	0.0047	0.122	0.122	0.122	0.049	0.122	0.122	0.122	0.049	0.049	0.122
5	0.0566	0.171	0.171	0.171	0.069	0.069	0.069	0.069	0.069	0.069	0.069
6	0.1114	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
7	0.0605	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
8	0.1112	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
9	-0.0248	0.156	0.156	0.156	0.063	0.063	0.063	0.063	0.063	0.063	0.156
10	0.1204	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
11	0.0585	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
12	0.1044	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100

Tabela 6. Resultados da Fase 2, com restrições aos coeficientes de ponderação e folgas livres (modelo 1).

Fase 2										
DMU	\hat{s}^*_{MCDT}	\hat{s}^*_{MED}	\hat{s}^*_{RH}	\hat{s}^*_{OC}	\hat{s}^*_{CPF}	\hat{s}^*_{CSM}	\hat{s}^*_{CSI}	\hat{s}^*_{CSA}	\hat{s}^*_{CSS}	\hat{s}^*_{SQ}
1	0.061	-0.037	0.213	0.201	-0.019	0.082	-0.078	-0.043	-0.051	-0.021
2	0.164	-0.002	-0.003	0.211	0.023	0.088	0.121	0.023	-0.018	0.329
3	0.167	-0.037	0.311	0.130	0.062	0.195	0.300	0.200	0.153	0.208
4	-0.053	0.052	-0.039	0.258	-0.103	-0.048	-0.053	0.031	-0.005	0.167
5	-0.127	-0.022	0.033	0.237	-0.007	0.064	0.231	0.049	0.019	0.509
6	0.283	0.381	0.372	0.385	-0.206	-0.159	-0.069	-0.110	-0.075	0.313
7	0.610	0.672	0.612	0.628	-0.392	-0.362	-0.445	-0.523	-0.473	0.279
8	0.339	0.337	0.565	0.809	-0.223	-0.413	-0.002	-0.308	-0.222	0.229
9										
10	0.277	0.469	0.186	0.355	-0.334	-0.084	-0.069	-0.110	-0.075	0.589
11	0.492	0.534	0.459	0.647	-0.594	-0.365	-0.259	-0.334	-0.258	0.263
12	0.344	0.440	0.566	0.493	-0.293	-0.338	-0.102	-0.211	-0.184	0.329

Neste caso há apenas uma unidade eficiente (DMU 9), que é a referência para todas as outras (Tabela 5). Esta unidade tem as melhores práticas de acordo com as preferências do GD.

Obteve-se a seguinte ordenação:

DMU 9 > DMU 4 > DMU 1 > DMU 5 > DMU 11 > DMU 7 > DMU 2 > DMU 12 > DMU 8 > DMU 6 > DMU 10 > DMU 3.

Ao GD agradou que a DMU 9 ficasse em primeiro lugar mas ficou um pouco surpreendido, pois esperava que outra unidade assumisse o primeiro lugar. Este resultado suscitou o debate no GD, tendo surgido como possível explicação o facto de os fatores de *output* traduzirem pouco o trabalho de enfermagem. A DMU que ficou em último lugar foi realmente surpreendente, mas o GD conjecturou que existe uma forte possibilidade de não terem sido registadas todas as consultas médicas naquela unidade. No entanto, seria necessário averiguar melhor a causa desta última posição neste modelo.

O modelo 2 é aquele que o GD considera como sendo o mais abrangente, dado que combina os vários eixos de avaliação correspondentes ao acesso, desempenho assistencial, qualidade percebida e desempenho económico e reflete todo o trabalho médico, de enfermagem e administrativo. Os dados originais e as utilidades correspondentes para o modelo 2 são expostos na Tabela 7

As restrições aos coeficientes de ponderação foram obtidas de forma semelhante às do modelo 1 e ao GD foi pedido para comparar os "swings" da utilidade de 0 para 1 apresentados na Tabela 8 .

A ordenação dos coeficientes de escala obtida foi:

$$w_{MED} \geq w_{MCDT} \geq w_{RH} \geq w_{DM} \geq w_{OC} \geq w_{SQ} \geq w_{CE} \geq w_{CM} \geq w_{DE}.$$

De acordo com a resposta conseguida à questão de julgamento de indiferença, foi estabelecida a seguinte desigualdade: $w_{MED} \leq 2.47 w_{DE}$.

A ordenação das unidades é (d^* mostrado entre parênteses):

DMU 1 (-0.0562) > DMU 9 (-0.0013) > DMU 4 (0.0229) > DMU 5 (0.0503) > DMU 2 (0.0725) > DMU 3 (0.1395) > DMU 6 (0.1668) > DMU 10 (0.1810) > DMU 11 (0.1886) > DMU 12 (0.2003) > DMU 7 (0.2429) > DMU 8 (0.2619).

Embora a DMU 9 não esteja na primeira posição, continua a ser eficiente neste modelo.

Tabela 7. Desempenhos das DMUs nas escalas originais e na escala de utilidades para o modelo 2.

Fatores nas escalas originais									
DMU	<i>X_{MCDT}</i>	<i>X_{MED}</i>	<i>X_{RH}</i>	<i>X_{OC}</i>	<i>Y_{DM}</i>	<i>Y_{CM}</i>	<i>Y_{DE}</i>	<i>Y_{CE}</i>	<i>Y_{SQ}</i>
1	362900	1114935	1398033	123382	985	48038	2006	24273	92
2	463549	1192330	1050558	128606	241	37977	1424	20580	74
3	466561	1113734	1589818	101750	96	8265	844	11677	81
4	277140	1321779	1001297	152974	308	38102	1979	17150	83
5	232565	1147295	1101200	141864	4	21665	792	27328	62
6	613535	2491923	1724028	218890	195	50768	3645	34610	75
7	1332317	4365605	2368401	360011	570	96530	4166	23781	77
8	701121	2289334	2225321	491140	182	68599	987	34052	80
9	314090	1196353	1054564	65416	302	40283	1756	17435	91
10	606148	2954054	1348651	203098	244	45819	2249	45119	54
11	1007909	3350321	1933791	371468	365	75904	5031	35040	78
12	70936	2793686	2227392	278936	246	64165	3878	34831	74
Fatores em escalas de utilidade									
DMU	<i>U_{MCDT}</i>	<i>U_{MED}</i>	<i>U_{RH}</i>	<i>U_{OC}</i>	<i>U_{DM}</i>	<i>U_{CM}</i>	<i>U_{DE}</i>	<i>U_{CE}</i>	<i>U_{SQ}</i>
1	0.633	0.824	0.572	0.743	0.731	0.194	0.185	0.225	0.750
2	0.529	0.789	0.788	0.733	0.120	0.147	0.108	0.170	0.400
3	0.527	0.824	0.475	0.815	0.047	0.008	0.032	0.039	0.521
4	0.746	0.735	0.825	0.687	0.153	0.148	0.181	0.120	0.563
5	0.820	0.809	0.753	0.708	0.001	0.071	0.025	0.274	0.220
6	0.411	0.406	0.413	0.560	0.097	0.207	0.441	0.404	0.417
7	0.084	0.115	0.173	0.317	0.308	0.478	0.528	0.217	0.450
8	0.355	0.450	0.220	0.136	0.090	0.304	0.051	0.394	0.500
9	0.694	0.787	0.785	0.945	0.150	0.158	0.152	0.124	0.729
10	0.416	0.318	0.599	0.590	0.121	0.184	0.217	0.642	0.140
11	0.202	0.253	0.326	0.298	0.182	0.349	0.672	0.411	0.467
12	0.350	0.347	0.219	0.452	0.122	0.276	0.480	0.408	0.400

Tabela 8. Desempenhos extremos associados com os níveis de utilidade 0 e 1 (modelo 2).

Nível de utilidade	X_{MCDT}	X_{MED}	X_{RH}	X_{OC}	Y_{DM}	Y_{CM}	Y_{DE}	Y_{CE}	Y_{SQ}
$u(.)=0$	1600000	5500000	3000000	600000	3	6500	600	9000	40
$u(.)=1$	150000	800000	800000	50000	1200	150000	6500	55000	100

Os resultados alcançados corresponderam às expectativas do GD, em vários aspectos: a DMU 3 recupera, subindo várias posições no *ranking*, com a inclusão do trabalho de enfermagem; a DMU 5 está na 2.^a posição entre as ineficientes, mantendo a 4.^a posição no *ranking* geral; a DMU 1 é, como esperado, a melhor de todas as unidades, sendo uma referência para a grande maioria das unidades ineficientes (exceto para uma); a DMU 8 é classificada como a mais ineficiente, facto nada surpreendente para o GD.

6. Melhorias das unidades ineficientes e outros desenvolvimentos

Considere-se que W denota o conjunto de vetores de coeficientes de ponderação compatíveis com o *ranking* e limite obtidos (ver 4.3.). A Fase 2 do método usa o vetor de ponderação ótimo em W^* para resolver o modelo de PL da Fase 2 com as variáveis folga a assumir qualquer valor real. A solução é uma proposta de um alvo de eficiência (projeção) para cada DMU ineficiente. Para atingir o estado de eficiência, essas DMUs ineficientes devem mudar a sua utilidade em cada fator pelo valor indicado por s^* , considerando os dois modelos.

À partida, não havia restrições aos alvos: qualquer alvo era admitido desde que permitisse à DMU anular o seu défice de valor global. No entanto, tal não impedia que o alvo escolhido fosse pior nalguns fatores, desde que estas variações fossem suficientemente compensadas por melhorias nos restantes fatores. Logo, nem todas as mudanças propostas correspondem a melhorias em todos os fatores, uma vez que algumas folgas podiam ser negativas. Note-se, porém, que uma folga negativa para um fator de *input* corresponde a uma sugestão de aumentar o consumo desse *input*, e uma folga para um fator de *output* corresponde a uma sugestão de reduzir a produção desse *output*. O GD não concordou com estas possibilidades, uma vez que não tinha vantagem

alguma em aumentar recursos, nomeadamente custos, nem em diminuir serviços prestados, pelo que houve a necessidade de propor uma outra solução.

Considera-se que u_k^* é a utilidade da melhor DMU usando o vetor de ponderação no conjunto ótimo W^* , ou seja, o valor da utilidade que a DMU k deve alcançar:

$$u_k^* = \sum_{c=1}^q w_c^* u_c(DMU_k) + d_k^* \quad (6)$$

Sabe-se que o modelo de PL resolvido na Fase 2 do método de Gouveia et al. (2008) admite alvos ótimos alternativos, cada um dos quais corresponde a uma forma diferente de anular o valor da diferença d_k^* . Esses alvos correspondem a diferentes projeções sobre a fronteira eficiente. No presente caso, o GD tem a intenção de restringir os alvos de eficiência que se propõem alcançar u_k^* , não só para evitar os alvos que implicam um aumento de *inputs* ou uma redução dos *outputs*, mas também para escolher, dadas as características de cada unidade, quais os fatores que podem ser alterados.

O conjunto $S_< = \{c \in \{1, \dots, q\}: s_c^* < 0 \text{ na Fase 2}\}$ denota as folgas negativas no ótimo de (5); essas folgas vão passar agora a ser constantes nulas. O conjunto $S_> = \{c \in \{1, \dots, q\}: s_c^* \geq 0 \text{ na Fase 2}\}$ denota as restantes folgas, que serão consideradas como variáveis não-negativas. Portanto, a formulação que conduz a um alvo alternativo que mantém ou melhora o desempenho de todos os fatores pode ser obtida resolvendo o modelo de PL (7), no qual a folga máxima (em termos de valor) deve ser minimizada para atingir o valor global de utilidade do alvo. Não são permitidas folgas negativas, mas o alvo já não será uma combinação convexa das DMUs observadas (para mais detalhes ver Almeida & Dias, 2012).

A formulação abaixo difere da utilizada por Almeida e Dias (2012) na medida em que introduz um novo parâmetro σ_c (um valor positivo) para limitar o valor que uma folga pode ter. O objetivo é evitar a fixação de alvos irrealistas, permitindo que as propostas de melhoria sejam mais ajustadas às exigências do GD. Os alvos nunca excedem o valor 1 em qualquer fator, devido à restrição $u_c(DMU_k) + s_c \leq 1$. Tal garante que a função de utilidade não ultrapassa os intervalos dos desempenhos obtidos.

$$\begin{aligned}
& \min_{\rho_k, s} \rho_k \\
& \text{s. a: } \sum_{c \in S_{\geq}} w_c^* s_c = d_k^* \\
& \quad u_c(DMU_k) + s_c \leq 1, \quad c = 1, \dots, q \\
& \quad s_c - \sigma_c \leq 0, \quad c \in S_{\geq} \\
& \quad \rho_c \geq s_c \geq 0, \quad c = 1, \dots, q
\end{aligned} \tag{7}$$

No que diz respeito ao modelo 1, todas as unidades ineficientes escolheram como par a DMU 9 e o GD começou por analisar os resultados em termos de propostas de melhoria das unidades ineficientes, bloqueando alterações nos fatores que tinham folgas negativas na Fase 2 (ver proposta 1 na Tabela 9), não permitindo que, nestes fatores, o alvo tivesse pior desempenho que o desempenho original das unidades.

Para algumas unidades a proposta 1 tem desempenhos atingíveis, mas para outras tais seriam impossíveis de alcançar. Esta realidade obrigou à construção de uma nova formulação, (8), na qual foi feita uma distribuição por todos os fatores (*inputs* e *outputs*) do valor da diferença relativamente ao par, de uma forma equilibrada, através do seguinte modelo de PL:

$$\begin{aligned}
& \min_{\rho_k, s} \rho_k \\
& \text{s. a: } \sum_{c=1}^q w_c^* s_c = d_k^* \\
& \quad u_c(DMU_k) + s_c \leq 1, \quad c = 1, \dots, q \\
& \quad s_c - \sigma_c \leq 0, \quad c = 1, \dots, q \\
& \quad \rho_c \geq s_c \geq 0, \quad c = 1, \dots, q
\end{aligned} \tag{8}$$

Mesmo assim, o GD não entendeu como praticáveis, dadas as características de certas unidades, algumas das reduções propostas, nomeadamente no custo total com recursos humanos proposto para a DMU 3, e nos aumentos de algumas consultas propostas para as DMUs 6 e 10 (ver proposta 2 na Tabela 9). Neste último caso a justificação está relacionada com estatísticas nacionais associadas à taxa de natalidade, bem como ao facto de alguns utentes (apesar de inscritos no serviço de saúde público) preferirem ir ao serviço de saúde privado, possivelmente por motivos de desempenho percebido. Assim, surgiu a proposta de limitar a redução ou aumento de alguns fatores de *input* ou *output*, para as unidades 3, 6 e 10 no modelo 1. Para este efeito, a diferença entre a unidade em avaliação e a DMU 9 deveria ser suprimida pela redução ou aumento de outros fatores de *input* ou *output* (ver proposta 3 na Tabela 9).

Tabela 9. Propostas de melhoria para as unidades ineficientes nas escalas originais (modelo 1)

Proposta 1										
DMU	X _{MCDT}	X _{MED}	X _{RH}	X _{OC}	Y _{CPF}	Y _{CSM}	Y _{CSI}	Y _{CSA}	Y _{CSS}	Y _{SQ}
1	311190	1114935	1285500	10415	845	462	5133	25380	17149	92
2	350315	1192330	1050558	85950	1741	610	4644	29710	14783	81.8
3	303800	1113734	1250900	50750	1780	472	3230	20340	15500	89.7
4	277140	1311200	1001297	145745	1628	605	4934	20422	13811	83.6
5	232565	1147295	970000	105000	732	573	3013	25570	19028	69
6	362000	1622900	1285000	110900	2534	782	5068	29422	18854	86.6
7	1000840	3462876	2009990	288090	3726	1081	8073	54200	40307	83.4
8	414500	1490900	1659300	335300	2641	1142	4530	41247	27858	90.6
9										
10	342900	1855000	981100	97400	3353	662	5068	29422	18854	72.8
11	763900	2672000	1658890	301300	4763	1084	6587	42838	30037	84.0
12	432500	1869000	1689900	166100	3087	1052	5328	35475	25593	85.2
Proposta 2										
DMU	X _{MCDT}	X _{MED}	X _{RH}	X _{OC}	Y _{CPF}	Y _{CSM}	Y _{CSI}	Y _{CSA}	Y _{CSS}	Y _{SQ}
1	344500	1068900	1358900	115650	1048	371	5308	26710	18740	92.7
2	381560	1020700	942490	101690	1207	488	4190	26300	14780	78.8
3	332600	846100	1316900	61750	1420	386	2680	17100	12700	87.8
4	274440	1288720	996200	150430	1666	612	4964	19667	14132	83.2
5	203699	1030610	1022900	115990	1254	486	2453	22522	18854	66.6
6	472200	2010000	1487990	160900	3242	960	5957	36100	25080	81.3
7	1153760	3887255	2187990	324100	4085	1153	8556	56760	43700	80.5
8	539200	1850900	1923300	402800	3351	1275	5415	47890	34490	85.3
9										
10	455000	2342000	1149100	140100	4090	855	6030	36610	26280	65.6
11	878540	2996000	1790990	336700	5040	1153	7052	46330	33580	81.2
12	553490	2287190	1939300	221100	3757	1177	6160	41710	31880	80.2
Proposta 3										
DMU	X _{MCDT}	X _{MED}	X _{RH}	X _{OC}	Y _{CPF}	Y _{CSM}	Y _{CSI}	Y _{CSA}	Y _{CSS}	Y _{SQ}
3	380500	1000900	1000000	101690	12175	489	4210	21340	10789	86.3
6	471200	2000000	1488010	160990	3249	800	5500	36190	25120	82.3
10	454890	2341090	1149090	140000	4150	800	5500	36623	26340	70

As propostas de melhoria para o modelo 1 em todos os cenários referidos são apresentadas na Tabela 9, com os fatores que não foram alterados ou estão limitados (proposta 3) mostrados em negrito.

O GD ficou muito satisfeito com a possibilidade de escolher, para cada caso, a proposta de melhoria que entende ser a mais adequada, uma vez que conhece bem as características específicas de cada unidade.

Para o modelo 2, e atendendo às unidades em avaliação e aos fatores em causa, o GD entendeu suficiente a apresentação apenas das propostas 1 e 2 (Tabela 10).

Tabela 10. Propostas de melhoria para as unidades ineficientes nas escalas originais (modelo 2).

Proposta 1									
DMU	X_{MCDT}	X_{MED}	X_{RH}	X_{OC}	Y_{DM}	Y_{CM}	Y_{DE}	Y_{CE}	Y_{SQ}
1									
2	376000	103900	1050558	99900	418	57000	2100	26480	79.3
3	298200	1113734	1240300	101750	470	48580	2280	24500	90
4	277140	1232500	1001297	134300	380	45000	2260	19600	84.7
5	232565	1147295	976900	106700	183	40850	1480	27328	68.6
6	339500	1535900	1238000	103300	616	50768	3645	34610	80
7	636500	2737000	1569100	183100	897	96530	4166	41090	92.3
8	316100	1195900	1426900	266500	712	68599	3328	34052	94.8
9									
10	319300	1753900	1348651	88700	672	92720	2249	45119	74.7
11	516690	19339800	1332000	208600	758	75904	5031	35040	89.6
12	347800	1566900	1498400	119580	707	64165	3878	34831	89.6
Proposta 2									
DMU	X_{MCDT}	X_{MED}	X_{RH}	X_{OC}	Y_{DM}	Y_{CM}	Y_{DE}	Y_{CE}	Y_{SQ}
1									
2	389790	1037990	953800	104500	384	53480	1972	25490	78.4
3	334900	851090	1323400	62470	372	38080	1910	21200	87.7
4	262985	126600	971500	141580	352	43000	2150	18690	84.1
5	206100	1042500	1031000	117900	105	32290	1180	30120	66.2
6	412800	1809000	1384000	132000	515	79760	4645	42560	84
7	749500	2392000	1719100	219000	841	126500	5585	37750	89.2
8	377000	13839000	1575400	311500	622	107860	2874	45598	92.4
9									
10	394800	2083000	1061700	114200	562	78320	3385	50740	69.2
11	644190	2332800	1507000	258600	645	104500	5938	43600	87.4
12	441590	1899900	1709000	171190	587	96100	5080	43880	84.8

7. Considerações Finais

Este estudo relata um processo interativo em que o analista e especialistas em saúde trabalharam juntos na avaliação de unidades de CSP, com o intuito de identificar as melhores práticas e contribuir para uma tomada de decisão informada. A abordagem DEA/MCDA proposta anteriormente pelos autores foi utilizada como base de modelação e posteriormente adaptada aos requisitos desta aplicação, ou seja, para definir alvos realistas (com a introdução de um novo parâmetro no modelo) no sentido de possibilitar a ordenação total das DMUs.

A intervenção para avaliar o desempenho das unidades de saúde começa com a escolha pelo GD dos fatores a serem incluídos no modelo, abrangendo quatro grandes áreas: acesso, desempenho assistencial, qualidade percebida e desempenho económico. Fruto das discussões, resultaram dois modelos para avaliar as unidades de CSP. A incorporação das preferências do GD nesses modelos foi possível graças à conversão dos fatores *inputs* e *outputs* em funções de utilidade, o que exigiu interpretar estas funções de utilidade como dispositivos para comparar a variação de mérito que corresponde a diferenças de desempenho. Os modelos tornam-se mais completos e credíveis para o GD apenas com a introdução de restrições de coeficientes de ponderação que foram obtidas recorrendo à técnica dos *swings* para estabelecer uma ordenação dos coeficientes de ponderação. Além disso, foi usada uma questão baseada em *trade-offs* para limitar a relação entre os coeficientes de ponderação posicionados no primeiro e último lugares e, desta forma, evitar coeficientes de ponderação nulos.

Para cada modelo obteve-se o *ranking* completo das unidades e foram identificados os pares que lhes serviriam de referência. As propostas de melhoria para as unidades ineficientes surgiram aproveitando o conhecimento do GD sobre as unidades, a fim de obter alvos mais realistas. Para este efeito, um novo modelo foi desenvolvido onde é introduzido um novo parâmetro que limita o valor que uma folga pode ter, possibilitando propostas de melhoria mais ajustadas às preferências do GD.

Este trabalho permitiu facultar informação ao GD no sentido da identificação das melhores práticas, fontes de ineficiência e diferenças em

relação às melhores práticas, bem como possibilidades e oportunidades de melhoria. Os resultados foram úteis não só para a conceção de medidas corretivas, mas também para informar as decisões sobre os objetivos dos IPs no futuro. Todo o processo de avaliação das unidades de CSP desenrolou-se alternando reuniões de discussão e apresentação de resultados com desenvolvimento metodológicos e testes computacionais, podendo este não ser um processo replicável em todos os casos. Os desenvolvimentos futuros deste trabalho incluem a implementação de um sistema de apoio à decisão que responda em tempo real aos requisitos dos utilizadores.

BIBLIOGRAFIA

ALI, A.I., & Seiford, L. (1993) The mathematical programming approach to efficiency analysis. In Fried, Lovell & Schmidt (Orgs.). *The measurement of productive efficiency: techniques and applications* (pp.120-113). New York: Oxford University.

ALI, A.I., Lerne, C.S., & Seiford, L. (1995). Components of efficiency evaluation in Data Envelopment Analysis. *European Journal of Operational Research*, 80, 462-473.

ALLEN, R., Athanassopoulos, A., Dyson, R.J., & Thanassoulis, E. (1997). Weights restrictions and value judgments in Data Envelopment Analysis: evolution, development and future directions. *Annals of Operations Research*, 73,13-34.

ALMEIDA, P.N., & Dias, L.C. (2012). Value-based DEA models: application-driven developments. *Journal of the Operational Research Society*, 63, 16-27.

AMADO, C.A.F., & Dyson, R.G. (2008). On comparing the performance of primary care providers. *European Journal of Operational Research*, 185, 915-932.

AMADO, C.A.F., & Santos, S.P. (2009). Challenges for performance assessment and improvement in primary health care: The case of the Portuguese health centres. *Health Policy*, 91, 43-56.

ANDERSEN, P., & Petersen, N.C. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science*, 39, 1261-1264.

ATHANASSOPOULOS, A.D., & Podinovski, V.V. (1997). Dominance and potential optimality multiple criteria models decision analysis with imprecise information. *Journal of the Operational Research Society*, 48 (1), 142-150.

BAKER, R. (1992). General practice in Gloucestershire, Avon and Somerset: Explaining variations in standards. *British Journal of General Practice*, 42, 415-418.

BANKER, R.D., Charnes, A., & Cooper, W.W. (1994). Some models estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management Science*, 3 (9), 1078-1092.

CAMPBELL, S.M., Roland, M.O., Quayle, J.A., Buetow, S.A., & Shekelle, P.G. (1998). Quality indicators for general practice: Which ones can general practitioners and health authority managers agree are important and how useful are they?. *Journal of Public Health Medicine*, 20, 414-421.

CHARNES, A., Cooper, W.W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2 (6), 429-444.

CHARNES, A., Cooper W.W., Golany, B., Seiford, L., & Stutz, J. (1995). Foundations of Data Envelopment Analysis for Pareto-Koopmans efficient empirical production functions. *Journal of Econometrics*, 30, 91-107.

Avaliação sobre o funcionamento das Unidades de Saúde do ACES – BM1. (2010). Coimbra.

Coodenação Estratégica dos CSP. (2010). *Metodologia de Contratualização*, Administração Central do Sistema de Saúde.

DIAS, L.C., & Clímaco, J.N. (2000). Additive Aggregation with Variable Interdependent Parameters: the VIP Analysis Software. *Journal of the Operational Research Society*, 51 (9), 1070-10822.

DONABEDIAN, A. (1980). *The Definition of Quality and Approaches to its Assessment*. Ann Arbor, MI, Health Administration Press.

DULÁ, J.H., & Hickman, B.L. (1997). Effects for excluding the column being scored from DEA envelopment LP technology matrix. *Journal of the Operational Research Society*, 48, 1001-1012.

DYSON, R.G., Allen, R., Camanho, A.S., Podinovski, V.V., Sarrico, C.S. & Shale, E.A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*, 132, 245-259.

FARRELL, M.J. (1997). The Measurement of productive efficiency. *Journal of Royal Statistical Society*, 120 (A), 353-290.

GIUFFRIDA, A., Gravelle, H., & Sutton, M. (2000). Efficiency and administrative costs in primary care. *Journal of Health Economics*, 19 (6), 983-1006.

GOODWIN, P., & Wright, G. (1998). *Decision analysis for management judgment*. Wiley: Chichester, 1998.

GOUVEIA, M.C., Dias, L.C., & Antunes, C.H. (2008). Additive DEA based on MCDA with imprecise information. *Journal of the Operational Research Society*, 59, 54-63.

GOUVEIA, M.C., Dias, L.C., & Antunes, C.H. (2013). Super-efficiency and stability intervals in additive DEA. *Journal of the Operational Research Society*, 64, 86-96.

HOLLINGSWORTH, B. (2003). Non-Parametric and Parametric Applications Measuring Efficiency in Health Care. *Health Care Management Science*, 6, 203-218.

HUANG, Y.G., & McLaughlin, C.P. (1989). Relative efficiency in rural primary health care: an application of data envelopment analysis. *Health Services Research*, 24, 143-158.

Instituto de Gestão Informática e Financeira da Saúde. (2006). *Estatística do Movimento Assistencial dos Centros de Saúde em 2005*.

KEENEY, R.L., Raiffa, H. (1976). *Decisions with multiple objectives: preferences and value tradeoff*. Wiley: New York.

KIRIGIA, J.M., Emrouznejad, A., Sambo, L.G., Munguti, N., & Liambila, W. (2004). Using data envelopment analysis to measure the technical efficiency of public health centers in Kenya. *Journal of Medical Systems*, 28, 155-166.

LINNA, M., Nordblad, A., & Koivu, M. (2003). Technical and cost efficiency of oral health care provision in Finnish health centres. *Social Science & Medicine*, 56, 343-353.

MILLIKEN, O., Devlin, R.A., Barham, V., Hogg, W., Dahrouge, S., & Russell, G. (2011). Comparative Efficiency Assessment of Primary Care Service Delivery Models Using Data Envelopment Analysis. *Canadian Public Policy*, 37, 85-109.

MCCOLL, A., Roderick, P., Gabbay, J., Smith, H., Moore, M. (1998). Performance indicators for primary care groups: An evidence based approach. *British Medical Journal*, 317, 1354-1360.

NUTI, S., Daraio, C., Speroni, C., & Vainieri, M. (2011). Relationships between technical efficiency and the quality and costs of health care in Italy. *International Journal for Quality in Health Care*, 23, 324-330.

Organização para a Cooperação e Desenvolvimento Económico. (2006). *Projecting OECD health and long term care expenditures: what are the main drivers?* Economics Department. Working Papers 477.

Observatório Português dos Sistemas de Saúde. (2007). *Relatório de Primavera 2007. Luzes e sombras: a governação da saúde*, Coimbra: Mar da Palavra.

Observatório Português dos Sistemas de Saúde. (2008). *Relatório de Primavera 2008. Sistema de Saúde Português – Riscos e Incertezas*, Coimbra: CEISUC.

OZCAN, Y.A. (1993). Sensitivity Analysis of Hospital Efficiency under Alternative Output/Input Combinations and Peer Groupings. *International Journal of Knowledge and Policy*, 4, 1-31.

PROCTOR, S., & Campbell, J. (1999). A developmental performance framework for primary care. *International Journal of Health Quality Assurance*, 12 (7), 279-286.

PODINOVSKI, V.V. (2004). Production trade-offs and weight restrictions in data envelopment analysis. *Journal of the Operational Research Society*, 55 (12), 1311-22.

ROSEMAN, R., & Friesner, D. (2004). Scope and scale inefficiencies in physician practices. *Health Economics*, 13, 1091-1116.

ROUSE, P., Harrison, J., & Turner, N. (2011). Cost and Performance: Complements for Improvement. *Journal of Medical Systems*, 35, 1063-1074.

SALINAS-JIMENEZ, J., & Smith, P. (1996). Data envelopment analysis applied to quality in primary health care. *Annals of Operational Research*, 67, 141-161.

SALO, A.A., & Hämäläinen, R.P. (2001). Preference Ratios in Multiattribute Evaluation (PRIME) - elicitation and decision procedures under incomplete information, *IEEE Transactions on Systems Man and Cybernetics part A*, 31, 533-545.

THANASSOULIS, E., Boussofiane, A., & Dyson, R.G. (1996). A comparison of data envelopment analysis and ratio analysis as tools for performance assessment. *Omega*, 24, 229-244.

THANASSOULIS, E., Portela, M.C., Allen, R. (2004). Incorporating value judgments in DEA. In Cooper, W.W., Seiford, L.M., Zhu, J. (eds.), *Handbook on Data Envelopment Analysis* (pp.99-137). Boston: Kluwer.

THANASSOULIS, E., Portela, M.C., Graveney, M. (2012). Estimating the scope for savings in referrals and drug prescription costs in the General Practice units of a UK Primary Care Trust. *European Journal of Operational Research*, 221, 432-444.

VON WINTERFELDT, D., Edwards, W. (1986). *Decision analysis behavioural research*. New York: Cambridge University Press.

VAN DEN HOMBERGH, P., Grol, R., van den Hoogen, H.J.M., & van den Bosch, W.J.H.M. (1998). Assessment of management in general practice: Validation of a practice visit method. *British Journal of General Practice*, 48, 1743-1750.

WAGNER, J., Shimshak, D., Novak, M. (2003). Advances in physician profiling: the use of DEA. *Socio-Economic Planning Sciences*, 37, 141-163.