



**HAL**  
open science

## Evaluation of the effects of the French Pay-for-Performance program - IFAQ pilot study

Benoît Lalloué, Shu Jiang, Anne Girault, Marie Ferrua, Philippe Loirat,  
Etienne Minvielle

► **To cite this version:**

Benoît Lalloué, Shu Jiang, Anne Girault, Marie Ferrua, Philippe Loirat, et al.. Evaluation of the effects of the French Pay-for-Performance program - IFAQ pilot study. *International Journal for Quality in Health Care*, 2017, 29 (6), pp.833-837. 10.1093/intqhc/mzx111 . hal-01579386

**HAL Id: hal-01579386**

**<https://hal.science/hal-01579386>**

Submitted on 31 Aug 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*This is a pre-copyedited, author-produced version of an article accepted for publication in International Journal for Quality in Healthcare following peer review. The version of record*

Benoît Lalloué, Shu Jiang, Anne Girault, Marie Ferrua, Philippe Loirat, Etienne Minvielle;  
Evaluation of the effects of the French pay-for-performance program—IFAQ pilot study, is  
available online at: <https://doi.org/10.1093/intqhc/mzx111> .

**Title:** Evaluation of the effects of the French Pay-for-Performance program - IFAQ pilot study

**Authors:**

Benoît Lalloué, PhD; EA 7348 MOS-EHESP, Gustave Roussy 114, rue Édouard-Vaillant  
94805 Villejuif, France ; [benoit.lalloue@gustaveroussy.fr](mailto:benoit.lalloue@gustaveroussy.fr)

Shu Jiang, PhD; EA 7348 MOS-EHESP, Gustave Roussy 114, rue Édouard-Vaillant 94805  
Villejuif, France ; [shu.jiang@u-psud.fr](mailto:shu.jiang@u-psud.fr)

Anne Girault, MSc; EA 7348 MOS-EHESP; [anne.girault@eleve.ehesp.fr](mailto:anne.girault@eleve.ehesp.fr)

Marie Ferrua, MSc; Gustave Roussy 114, rue Édouard-Vaillant 94805 Villejuif, France ;  
[marie.ferrua@gustaveroussy.fr](mailto:marie.ferrua@gustaveroussy.fr)

Philippe Loirat, MD; EA 7348 MOS-EHESP, Gustave Roussy 114, rue Édouard-Vaillant  
94805 Villejuif, France ; [loirat.philippe@wanadoo.fr](mailto:loirat.philippe@wanadoo.fr)

Etienne Minvielle, PhD, MD; EA 7348 MOS-EHESP, Gustave Roussy 114, rue Édouard-  
Vaillant 94805 Villejuif, France ; [etienne.minvielle@gustaveroussy.fr](mailto:etienne.minvielle@gustaveroussy.fr)

**Corresponding author:**

Benoît Lalloué  
EA 7348 MOS-EHESP Gustave Roussy  
114, rue Édouard-Vaillant  
94805 Villejuif Cedex  
FRANCE  
[benoit.lalloue@gustaveroussy.fr](mailto:benoit.lalloue@gustaveroussy.fr)

Tel: +33 (0)1 42 11 62 21

Fax: +33 (0)1 42 11 64 53

**Title:** Evaluation of the effects of the French Pay-for-Performance program - IFAQ pilot study

**Running title:** Impact of a French P4P program

**Abstract word count:** 250

**Text word count:** 1401

## **Abstract**

*Objective:* Most studies showed no or little effect of Pay-for-Performance (P4P) programs on different outcomes. In France, the P4P program IFAQ was generalized to all acute care hospitals in 2016. A pilot study was launched in 2012 to design, implement and assess this program. This article aims to assess the immediate impact of the 2012-2014 pilot study.

*Design and setting:* From 9 process QIs, an aggregated score was constructed as the weighted average, taking into account both achievement and improvement. Among 426 eligible volunteer hospitals, 222 were selected to participate. Eligibility depended on documentation of quality indicators and results of hospital accreditation. Hospitals with scores above the median received a financial reward based on their ranking and budget. Several characteristics known to have an influence on P4P results (patient age, socioeconomic status, hospital activity, casemix, and location) were used to adjust the models.

*Intervention:* To assess the effect of the program, comparison between the 185 eligible selected hospitals and the 192 eligible not selected volunteers were done using the difference-in-differences method.

*Results:* Whereas all hospitals improved from 2012 to 2014, the difference-in-differences effect was positive but not significant both in the crude (2.89,  $p=0.29$ ) and adjusted models (4.07,  $p=0.12$ ).

*Conclusion:* These results could be explained by several reasons: low level of financial incentives, unattainable goals, too short study period. However, the lack of impact for the first year should not undermine the implementation of other P4P programs. Indeed, the pilot study helped to improve the final model used for generalization.

**Keywords:** Pay-for-performance, hospital, impact evaluation

## **Introduction**

In many countries, healthcare systems are moving toward new payment models for hospitals (1). Among them, pay-for-performance (P4P) programs aimed at improving quality of care have been under high scrutiny. At the same time, concerns about their effectiveness have been raised.

The implementation and evaluation of P4P programs have been studied extensively, particularly the Medicare and Medicaid Hospital Value-Based Purchasing program in the United States. As for now, most studies have showed no or little effect of P4P on a range of outcomes (2–5).

In France, the P4P program IFAQ (*Incitation financière à l'amélioration de la qualité* - Financial Incentive to Quality Improvement) coordinated by the French Ministry of Health and the National Authority for Health (*Haute Autorité de Santé* - HAS) was generalized to all acute care hospitals in 2016 (about 1300 hospitals).

Before implementing the program nationally, it was first designed and tested in an experimental setting in 2012 within 222 hospitals and then assessed in 2014.

This article aims to assess the immediate impact of the 2012-2014 pilot study.

## **Material & Methods**

The IFAQ program was designed by a group of researchers, experts and representatives of French hospital federations under the authority of the Ministry of Health and HAS.

Of the 426 eligible volunteer hospitals, 222 were selected (170 randomly selected after stratification by type and administrative regions and 52 chosen directly by the hospital federations). Hospitals with undocumented quality indicators (QIs) or hospitals that were only conditionally accredited by HAS during the course of the pilot study were subsequently excluded from the sample, thus reducing the number to 185 hospitals that could be eligible to the payment of a bonus.

In each hospital, for each of 9 available process QIs, both achievement and improvement were taken into account to create a 0 to 10 score (tables 1 and 2). These QIs have been mandatory and publicly available since 2006 (6). Various domains were measured: care quality and safety, accreditation, electronic medical record. An aggregated score was constructed as the weighted average of these values on a 0 to 100 scale. The weights were allocated by the working group and ranged from 3.4% to 15% of the final score.

In this study, two periods were compared: the 2009-2011 period, before the intervention (achievement in 2011 and improvement between 2009 and 2011), and the 2011-2013 period, during

the intervention (achievement in 2013 and improvement between 2011 and 2013). Since two of the 9 QIs (electronic medical records and accreditation focus priority standards) did not have enough anteriority to compute the score during the pre-intervention period (no 2009 data), a score based on the 7 remaining QIs was finally calculated.

In the 2012-2014 pilot study, hospitals with scores above the median received a 3-tiered financial reward, based on their ranking and budget. They could receive from 0.3% to 0.5% of their annual budget with minimum and maximum payments of 50 k€ (56 k\$) and 500 k€ (560 k\$), respectively. Details about IFAQ are published elsewhere (7–10) and are synthesized in Appendix 1.

For each hospital, several characteristics, known to have an influence on P4P results (11,12) (volume of activity; DRGs accounting for 80% stays; number of authorized activities; percentage of patient stays <4 years old, 75-85 years old or >85 years old; percentage of low income patients; geographical location), were retrieved from medical and administrative databases (the PMSI, a mandatory national database which contains diagnosis and acts codes of all patients' hospital stays; the SAE, a mandatory and exhaustive survey among all hospitals; and the INSEE databases, national institute of statistics and economic studies).

To evaluate the impact, we compared the 185 hospitals (namely the “IFAQ group”) with the 192 eligible volunteers which were not selected in the final sample (namely “the control group”) using the difference-in-differences method, which is commonly used for this purpose (13).

This method allows to take into account the “natural trend” of the selected indicators (to estimate the trend of the aggregated score). Under the hypothesis that this trend is the same in both groups, the difference between the progression slopes of the control and IFAQ groups accounts for the effect of IFAQ. Difference-in differences effect can be assessed with regression models, which permit to adjust for explanatory variables. We used both crude and characteristics-adjusted OLS regression models to assess the impact of the program, using the score and its QIs components as dependent variables and the intervention group, the time period and the interaction between these two variables (which is the difference in difference effect) as independent variables. Hospital characteristics were added as supplementary independent variables for the characteristics-adjusted model.

The analysis was performed using the R 3.3.0 software (14).

## **Results**

The 2009-2011 mean score was 48.3 and 44.3 for the control and IFAQ groups, respectively. Both groups progressed during the study period, with a 2011-2013 mean score of 58.6 for the control group (10.3 points of increase) and 57.4 for the IFAQ group (13.1 points of increase) (table 3 and appendix 2).

As for the aggregated score, the difference-in-differences effect was positive but not significant in the crude model (2.89, 95%CI [-2.44; 8.22],  $p=0.29$ ). The effect increased when hospitals characteristics were taken into account in the adjusted model but was still not significant (4.07, 95%CI [-1.04; 9.17],  $p=0.12$ ) (table 4). Similar results were obtained for QIs' scores: no significant effects were shown by crude models and only a small effect on pain assessment traceability was detected by adjusted models (0.97, 95%CI [0.05; 1.89],  $p=0.04$ ).

## **Discussion**

While all hospitals improved between the 2009-2011 period and the 2011-2013 period, we found no significant impact of the incentive on the score or its components. These findings seem consistent with those of other international studies (2–5).

Many reasons could explain this result. Firstly, implanting a quality incentive program in a context of general improvement makes the impact evaluation complex, especially when the sample is relatively small.

Secondly, it could be due to a low level of financial incentives and hospitals could have thought that the potential reward was not worth the efforts invested to improve their score (15).

Thirdly, being awarded a bonus or not may rely on inappropriate or unattainable goals. For instance, decreasing mortality is a desired goal but a financial incentive based on improvement of quality indicators may be unable to trigger quality improvements actions aimed at decreasing mortality. (16). In our study, process indicators were selected on purpose by the working group, as they appeared appropriate to stimulate continuous quality improvement within hospitals (17).

Fourthly, 2013 data that were used to measure achievement were collected when guidelines on program implementation could have been insufficiently assimilated by all participants. Considering the complexity of the program, several years are probably necessary before it can be accepted by the staff and put into practice. A longer study period may allow better observation of long term outcomes(5). Indeed, preliminary qualitative results on the implementation of IFAQ among hospitals showed that information dissemination within hospitals remained limited to the executive level, leaving physicians, nurses, medical assistants and other professionals out of the scope of the

program despite their implication in quality improvement on a day-to-day basis and their specific involvement in the incentivized processes of care.

Fifthly, participation to the program was made among voluntary hospitals. Therefore, some hospitals may already have initiated quality improvement projects, even if they were not selected in IFAQ. Moreover, this type of initiative produces “spillover effects” between the two groups: hospitals may have anticipated that P4P will be implemented in the near future, thus initiating quality improvement projects, which would lead to an improvement in the score (18).

Finally, the difference-in-difference method may be too simple to detect an impact. Other methods, such as measuring local average treatment effect (LATE) or the quantile treatment effect on the treated (QTET) could be used in the next impact studies. Indeed, absence of an average effect, as measured by the difference-in-differences method, does not mean that the program had no effect at all. Identifying a positive effect on hospitals with the lowest scores only could justify maintaining the program, as it is a way to encourage them to improve their score and reduce quality gaps between hospitals.

In conclusion, the lack of impact observed for the first year after implementation should not undermine the development of P4P programs in the future, especially if the purpose is to reduce inequalities between hospitals. Moreover, this pilot study, as well as a second pilot study conducted in 2015 on an extended sample of about 500 hospitals were useful because they have provided policy makers with important information when developing the final model of the French hospital P4P that is used for all French acute care hospitals since 2016.

**Acknowledgements:** The authors want to thank the French Ministry of Health; the National Authority for Health; and all the members of the IFAQ working group for their help during the IFAQ project. The authors also want to thank the two reviewers for their very constructive review of this article.

**Funding:** The work was supported by the French Ministry of Health and the National Authority for Health. They provided some of the data but had no role in its analysis and interpretation, or in the preparation, review or approval of the manuscript.

## References:

1. Groene O, Skau JKH, Frølich A. An international review of projects on hospital performance assessment. *Int J Qual Health Care*. 2008 Jan 6;20(3):162–71.
2. Figueroa JF, Tsugawa Y, Zheng J, Orav EJ, Jha AK. Association between the Value-Based Purchasing pay for performance program and patient mortality in US hospitals: observational study. *BMJ*. 2016 May 9;353:i2214.
3. Ryan AM, Krinsky S, Kontopantelis E, Doran T. Long-term evidence for the effect of pay-for-performance in primary care on mortality in the UK: a population study. *The Lancet* [Internet]. 2016 May [cited 2016 May 24]; Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0140673616002762>
4. Markovitz AA, Ryan AM. Pay-for-Performance: Disappointing Results or Masked Heterogeneity? *Med Care Res Rev* [Internet]. 2016 Jan 7 [cited 2016 Mar 10]; Available from: <http://mcr.sagepub.com/cgi/doi/10.1177/1077558715619282>
5. Chee TT, Ryan AM, Wasfy JH, Borden WB. Current State of Value-Based Purchasing Programs. *Circulation*. 2016 May 31;133(22):2197–205.
6. Couralet M, Leleu H, Capuano F, Marcotte L, Nitenberg G, Sicotte C, et al. Method for developing national quality indicators based on manual data extraction from medical records. *BMJ Qual Saf*. 2013 Feb;22(2):155–62.
7. Jiang S, Couralet M, Girault A, Fourcade A, LeVaillant M, Loirat P, et al. The Rationale for the French Hospital Experiment with P4P (IFAQ): Lessons from abroad. *J Gest Déconomie Médicales*. 2013 Apr 1;30(7):435–53.
8. Santé.gouv.fr. Incitation financière à l'amélioration de la qualité (IFAQ) [Internet]. 2014 [cited 2014 Sep 9]. Available from: <http://www.sante.gouv.fr/incitation-financiere-a-l-amelioration-de-la-qualite-ifaq,11754.html>
9. Ferrua M, Fourcade A, Lalloué B, Girault A, Jiang S, Loirat P, et al. Incitation Financière à l'Amélioration de la Qualité (IFAQ) pour les établissements de santé français : Résultats de l'expérimentation (2012-2014). *J Gest Déconomie Médicales*. 2015;33(4–5):277–90.
10. Dozol A, Fouchard A, Lalloué B, Poulain X, Grenier C. Isqua16-3191 What Can Be Learned from the Implementation of a Pay for Performance Programme? *Int J Qual Health Care*. 2016 Oct 14;28(suppl 1):67–8.
11. Figueroa JF, Wang DE, Jha AK. Characteristics of hospitals receiving the largest penalties by US pay-for-performance programmes. *BMJ Qual Saf*. 2016 Mar 3;bmjqs-2015-005040.
12. Werner RM, Kolstad JT, Stuart EA, Polsky D. The Effect Of Pay-For-Performance In Hospitals: Lessons For Quality Improvement. *Health Aff (Millwood)*. 2011 Apr 6;30(4):690–8.
13. Dimick J, Ryan A. Methods for evaluating changes in health care policy: The difference-in-differences approach. *JAMA*. 2014 décembre;312(22):2401–2.
14. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2016. Available from: <https://www.R-project.org>



15. Jha AK. Time to get serious about pay for performance. *JAMA*. 2013 Jan 23;309(4):347–8.
16. Loirat P, Ferrua M, Lalloué B, Fourcade A, Minvielle E. Should payment for performance depend on mortality? *BMJ*. 2016 Jun 22;i3429.
17. Yang JH, Kim SM, Han SJ, Knaak M, Yang GH, Lee KD, et al. The impact of Value Incentive Program (VIP) on the quality of hospital care for acute stroke in Korea. *Int J Qual Health Care*. 2016 Oct 10;28(5):580–5.
18. Eijkenaar F, Emmert M, Scheppach M, Schöffski O. Effects of pay for performance in health care: a systematic review of systematic reviews. *Health Policy Amst Neth*. 2013 May;110(2–3):115–30.

**Table 1. Quality indicators used to calculate the IFAQ score**

Quality indicator	Description	Way to include it in IFAQ	
		Achievement	Improvement
Traceability of pain assessment <sup>1</sup>	Rate of patient records for which a pain assessment was performed	✓	✓
Quality and content of the medical record <sup>1</sup>	Score assessing the presence of up to 10 regulatory criteria in the hospitalized patient file.	✓	✓
Quality and content of the anaesthesia record <sup>1</sup>	Score assessing the presence of up to 13 criteria in the anaesthetic file.	✓	✓
Quality and content of multidisciplinary meetings in oncology <sup>1</sup>	Score assessing records of a dated multidisciplinary consultation meeting, including the treatment proposal and conducted with at least three health professionals from different specializations during the initial management of a first cancer diagnosis.	✓	✓
Time elapsed before sending discharge letters <sup>1</sup>	Rate of patient records for which the hospital discharge letter was sent out within eight days or less (regulatory time) and contained the qualitative information essential for continuity of care	✓	✓
Screening for nutritional disorders <sup>1</sup>	Rate of patient records for which a weight measurement, BMI calculation and weight evolution before admission were performed and indicated in the patient file within the first two days of the stay in adult patients	✓	✓
Focus priority topics standards <sup>2</sup>	Score assessing compliance to 8 priority standards in the accreditation procedure	✓	
Composite index for evaluation of activities against nosocomial infections <sup>3</sup>	Score assessing the organization of hospital-acquired infection control on the basis of 55 criteria relating to the organization, resources and actions of the healthcare facility.	✓	
Medical record digitization <sup>2</sup>	Score assessing the development level of electronic medical record	✓	

<sup>1</sup> These QIs are publicly available in a letter format: from A for the best hospitals to C for the least performing. These letters take into account the fact that QIs are computed from a sample of patients' medical records.

<sup>2</sup> These QIs are not directly available in a letter format and were converted using rules established by the IFAQ working group: from A for the best hospitals to C for the least performing.

<sup>3</sup> This QIs is publicly available in a letter format: from A for the best hospitals to E for the least performing. In IFAQ, letters C, D, E were considered equally "bad".

**Table 2.** QIs' score computation rules

Year « n-2 » score <sup>1</sup>	Year « n » score	Component score (taking into account both achievement and improvement)
A <sup>2</sup>	A	10
A	B	3.8
A	C, D, E	1.0
B	A	7.8
B	B	5.0
B	C, D, E	2.2
C, D, E	A	9.0
C, D, E	B	6.2
C, D, E	C	0.0
Not answered <sup>3</sup>	A, B, C, D, E	0.0
Not eligible <sup>4</sup>	A	10
Not eligible	B	5
Not eligible	C, D, E	0

<sup>1</sup> QIs used in the IFAQ program are computed every 2 years, therefore improvement is the evolution between results in one year and 2 years before

<sup>2</sup> Most of French QIs are computed from a sample of patient files and their resulted displayed in 3 or 5 letters to take into account confidence interval

<sup>3</sup> All QIs used in IFAQ are mandatory but some hospital did not answered anyway

<sup>4</sup> All QIs used in IFAQ need a minimum amount of hospital stays to be computed, therefore some hospitals were not eligible to some QIs

**Table 3.** Mean (sd) and evolution of the score components (by group and time period) and of the hospital characteristics

	Control group			Intervention group		
	2009-2011	2011-2013	Evolution	2009-2011	2011-2013	Evolution
IFAQ score <sup>1</sup>	48.3 (18.9)	58.6 (17.9)	10.3	44.3 (20.5)	57.4 (16.6)	13.1
Quality and content of the medical record <sup>2</sup>	4.8 (3.8)	5.6 (3.6)	0.8	4.5 (3.9)	5.7 (3.5)	1.2
Time elapsed before sending discharge letters <sup>2</sup>	1.2 (2.7)	1.7 (2.9)	0.5	1.4 (2.9)	1.9 (3.1)	0.5
Traceability of pain assessment <sup>2</sup>	6.4 (3.4)	7.4 (2.9)	0.9	5.1 (4.2)	7.0 (3.2)	1.9
Screening for nutritional disorders <sup>2</sup>	1.2 (2.8)	2.2 (3.5)	0.9	1.1 (2.8)	1.9 (3.3)	0.8
Quality and content of the anesthesia record <sup>2</sup>	7.0 (3.3)	7.8 (3.0)	0.8	6.1 (4.0)	7.6 (3.0)	1.5
Quality and content of multidisciplinary meetings in oncology <sup>2</sup>	4.6 (3.4)	5.2 (2.7)	0.6	4.9 (3.5)	5.7 (2.9)	0.8
Composite index for evaluation of activities against nosocomial infections <sup>3</sup>	6.7 (4.1)	9.0 (2.6)	2.3	6.5 (4.1)	8.5 (3.2)	1.9
Number of DRGs accounting for 80% of stays		58.6 (35.5)			65.6 (40.3)	
Volume of activities (in stays)		14,935 (11,013)			22,176 (24,788)	
Percent of patients stays <4 years old		4.7% (5.1%)			5.3% (6.2%)	
Percent of patients stays 75-85 years old		17.9% (6.7%)			18.0% (7.5%)	
Percent of patients stays >85 years old		7.1% (8.0%)			8.6% (9.7%)	
Number of authorized activities		9.7 (4.0)			10.9 (5.2)	
Percent of low income patients		15.0% (11.8%)			16.2% (11.6%)	
Geographical location (n (%)):						
<i>Large cities with suburbs</i>		170 (90.4%)			154 (87.0%)	
<i>Mid-sized and small cities with suburbs</i>		18 (9.6%)			20 (11.3%)	
<i>Isolated townships</i>		0 (0%)			3 (1.7%)	

<sup>1</sup> Score between 0 and 100, weighted sum of the 7 QIs<sup>2</sup> Score between 0 and 10, taking into account improvement and achievement on this QI<sup>3</sup> Score between 0 and 10, taking into account achievement only

**Table 4.** Difference in differences effects of the 2014 IFAQ pilot study on the score and its components

	Weight in the score <sup>1</sup>	Crude model			Adjusted model <sup>2</sup>		
		Coeff	95% CI	p	Coeff	95% CI	p
IFAQ score <sup>3</sup>	-	2.89	(-2.44; 8.22)	0.29	4.07	(-1.04; 9.17)	0.12
Quality and content of the medical record <sup>4</sup>	15.3	0.36	(-0.71; 1.43)	0.51	0.44	(-0.63; 1.50)	0.42
Time elapsed before sending discharge letters <sup>4</sup>	13.2	-0.09	(-0.93; 0.75)	0.84	0.03	(-0.79; 0.86)	0.93
Traceability of pain assessment <sup>4</sup>	18.3	0.97	(-0.03; 1.97)	0.06	0.97	(0.05; 1.89)	0.04
Screening for nutritional disorders <sup>4</sup>	10.2	-0.17	(-1.07; 0.74)	0.72	-0.01	(-0.88; 0.85)	0.97
Quality and content of the anesthesia record <sup>4</sup>	14.3	0.73	(-0.28; 1.74)	0.16	0.89	(-0.14; 1.91)	0.09
Quality and content of multidisciplinary meetings in oncology <sup>4</sup>	14.3	0.19	(-0.86; 1.23)	0.73	0.15	(-0.93; 1.23)	0.79
Composite index for evaluation of activities against nosocomial infections <sup>5</sup>	14.3	-0.32	(-1.47; 0.83)	0.59	-0.41	(-1.58; 0.77)	0.50

<sup>1</sup> Weights were allocated by the working group

<sup>2</sup> Adjusted on volume of activity; DRGs accounting for 80% stays; authorized activities; percentage of patient stays <4 years old, 75-85 years old, >85 years old; percentage of low income patients; geographical location

<sup>3</sup> Score between 0 and 100, weighted sum of the 7 QIs

<sup>4</sup> Score between 0 and 10, taking into account improvement and achievement on this QI

<sup>5</sup> Score between 0 and 10, taking into account achievement only

## **Appendix 1. Details about IFAQ**

The following tables describe the different computation steps of the IFAQ score and reward:

- Step 1: Listing the quality indicators used to calculate the IFAQ score
- Step 2: Creating a score by QI, taking into account both achievement and improvement
- Step 3: Computing the IFAQ score by weighted average of the QI scores
- Step 4: Computing the reward

<b>Step 1. Quality indicators used to calculate the IFAQ score</b>			
<b>Quality indicator</b>	<b>Description</b>	<b>Way to include it in IFAQ</b>	
		<b>Achievement</b>	<b>Improvement</b>
Traceability of pain assessment*	Rate of patient records for which a pain assessment was performed	✓	✓
Quality and content of the medical record*	Score assessing the presence of up to 10 regulatory criteria in the hospitalized patient file.	✓	✓
Quality and content of the anaesthesia record*	Score assessing the presence of up to 13 criteria in the anaesthetic file.	✓	✓
Quality and content of multidisciplinary meetings in oncology*	Score assessing records of a dated multidisciplinary consultation meeting, including the treatment proposal and conducted with at least three health professionals from different specializations during the initial management of a first cancer diagnosis.	✓	✓
Time elapsed before sending discharge letters*	Rate of patient records for which the hospital discharge letter was sent out within eight days or less (regulatory time) and contained the qualitative information essential for continuity of care	✓	✓
Screening for nutritional disorders*	Rate of patient records for which a weight measurement, BMI calculation and weight evolution before admission were performed and indicated in the patient file within the first two days of the stay in adult patients	✓	✓
Focus priority topics standards**	Score assessing compliance to 8 priority standards in the accreditation procedure	✓	
Composite index for evaluation of activities against nosocomial infections***	Score assessing the organization of hospital-acquired infection control on the basis of 55 criteria relating to the organization, resources and actions of the healthcare facility.	✓	
Medical record digitization**	Score assessing the development level of electronic medical record	✓	
<p>* These QIs are publicly available in a letter format: from A for the best hospitals to C for the least performing. These letters take into account the fact that QIs are computed from a sample of patients' medical records.</p> <p>** These QIs are not directly available in a letter format and were converted using rules established by the IFAQ working group.</p> <p>*** This QIs is publicly available in a letter format: from 1 for the best hospitals to E for the least performing.</p>			

<b>Step 2. Creation of a score by QI, taking into account achievement and improvement</b>		
Year « n-2 » score	Year « n » score	Component score
A	A	10
A	B	3.8
A	C, D, E	1.0
B	A	7.8
B	B	5.0
B	C, D, E	2.2
C, D, E	A	9.0
C, D, E	B	6.2
C, D, E	C	0.0
Not answered*	A, B, C, D, E	0.0
Not eligible**	A	10
Not eligible**	B	5
Not eligible**	C, D, E	0

\* All QIs used in IFAQ are mandatory but some hospital did not answered anyway  
\*\* All QIs used in IFAQ need a minimum amount of hospital stays to be computed, therefore some hospitals were not eligible

<b>Step 3. Computation of the IFAQ score by weighted average of the QIs' scores</b>	
Quality indicator	Weight*
Traceability of pain assessment	15.1
Quality and content of the medical record	12.6
Quality and content of the anaesthesia record	11.8
Quality and content of multidisciplinary meetings in oncology	11.8
Time elapsed before sending discharge letters	10.9
Screening for nutritional disorders	8.4
Focus priority topics standards	14.3
Composite index for evaluation of activities against nosocomial infections	11.8
Medical record digitization	3.4

\*When all QIs applicable. If not, the same relative weights are kept.



<b>Step 4. Computation of the reward</b>				
<b>Hospital's rank</b>	<b>Tier</b>	<b>% of the annual budget</b>	<b>Bottom reward</b>	<b>Reward ceiling</b>
1 <sup>st</sup> to 31 <sup>st</sup>	0.5%-reward	0.5%	50k€ (56k\$)	500k€ (560k\$)
32 <sup>nd</sup> to 62 <sup>nd</sup>	0.4%-reward	0.4%	50k€ (56k\$)	400k€ (448k\$)
63 <sup>rd</sup> to 93 <sup>rd</sup> (median score)	0.3%-reward	0.3%	50k€ (56k\$)	300k€ (336k\$)
93 <sup>rd</sup> to the last	No reward	-	-	-

Appendix 2.

### IFAQ score and groups average adjusted-evolutions

(NB: a slight random noise was added on the x axis to give a better visualization of the data)

