atih

AGENCE TECHNIQUE
DE L'INFORMATION
SUR L'HOSPITALISATION

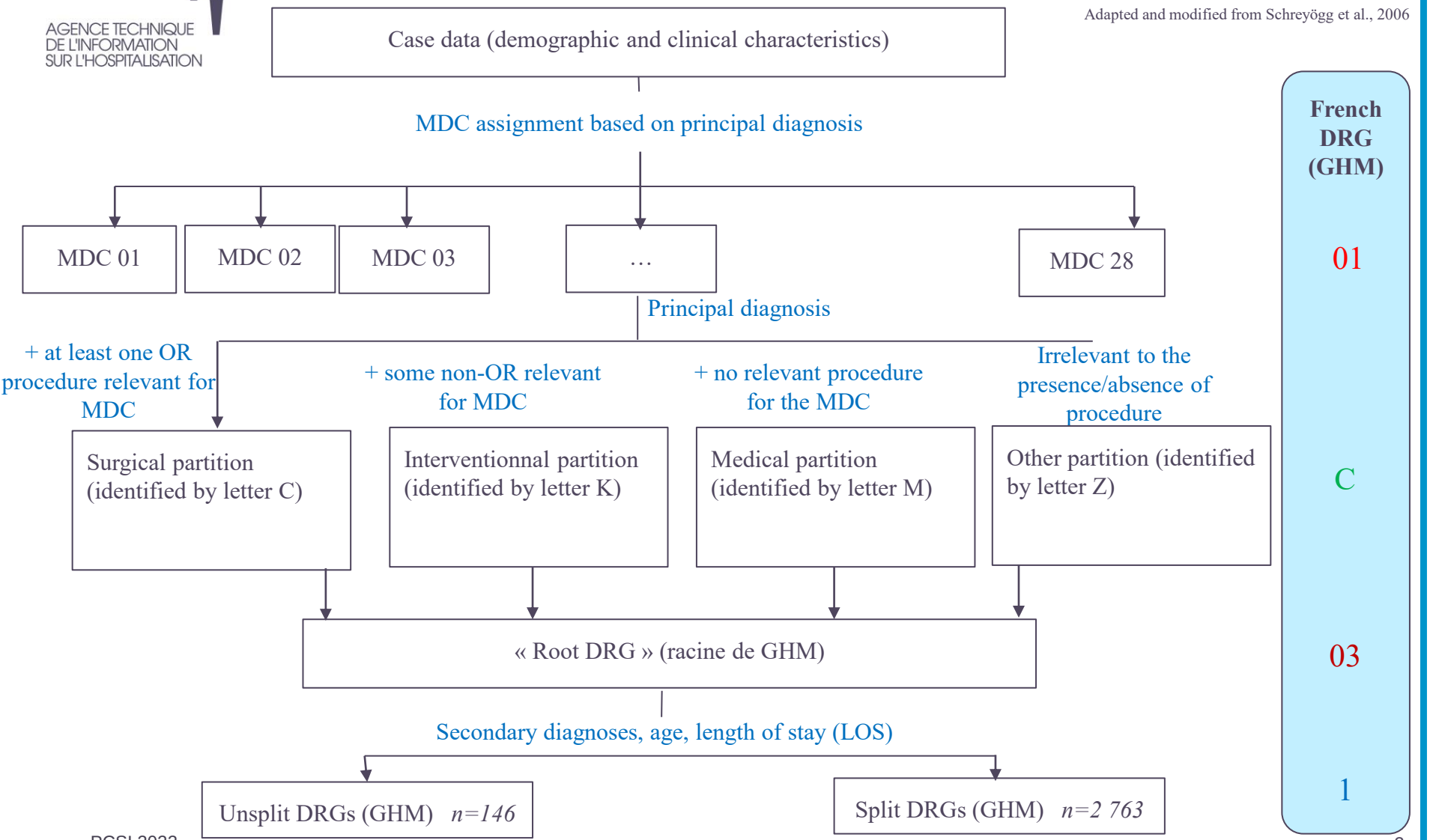
How to improve severity determination in the French DRGs

Nicolas Dapzol
Vincent Pisetta
Raphaël Schwob
Alexandra Delannoy
Raphaël Simon

117, bd Marius Vivier Merle
69329 Lyon cedex 03
Téléphone : 04 37 91 33 10
Fax : 04 37 91 33 67
www.atih.sante.fr

Overview of the French DRGs in acute care

Adapted and modified from Schreyögg et al., 2006



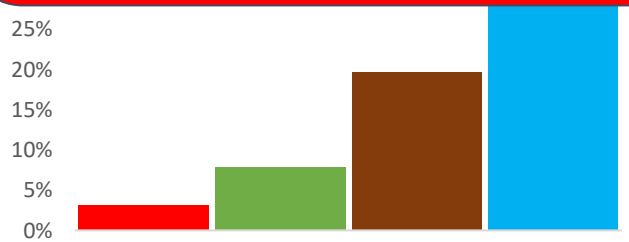
Severity index: Importance and evolution

R^2(GHM,LOS) in 2019

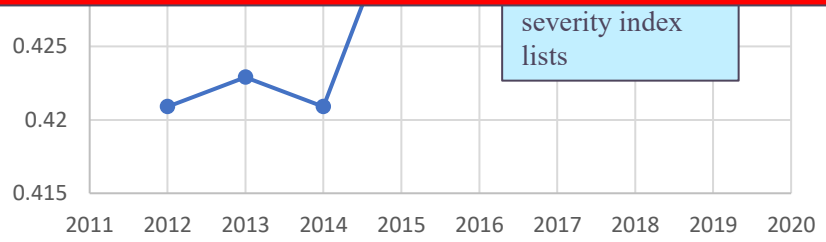
Evolution of R^2(LOS,GHM) over last years

01 C 03 1

MOTIVATION FOR UPDATING THE WAY SEVERITY IS COMPUTED



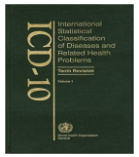
Severity index is essential in the quality of French DRG



Explicative power of severity index seems to reach a plateau

Determination of the Severity index in French DRGs

atih
AGENCE TECHNIQUE DE L'INFORMATION SUR L'HOSPITALISATION



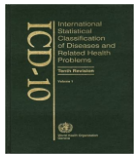
	STEP 1. Remove secondary diagnoses (SD) that are related to the principal diagnosis (PD) (*exclusion mechanism*)	STEP 2. The severity index of the stay is the max level of its SD	STEP 3. The severity index of the stay is potentially increased by 1 depending on the root-DRG and the age of the patient
PD	N39.0	N39.0	N39.0
Root DRG	11M04	11M04	11M04
Age	82	82	82
LOS	4	4	4
Secondary diagnoses (SD)	~~B96.2~~ C78.0 Z29.0 U82.2 A41.5	~~B96.2~~ C78.0 1 Z29.0 3 U82.2 3 A41.5 4	~~B96.2~~ C78.0 1 Z29.0 3 U82.2 3 A41.5 4

STEP 4. The severity index of the stay is potentially decreased if the LOS is lower than a specified threshold:
3 days min. for lvl 2
4 days min. for lvl 3
5 days min. for lvl 4

→ Level 3

Determination of the Severity index in French DRGs

AGENCE TECHNIQUE DE L'INFORMATION SUR L'HOSPITALISATION



	STEP 1. Remove secondary diagnosis (SD) that are related to the principal diagnosis (PD) (*exclusion mechanism*)	STEP 2. The severity index of the stay is the max level of its SD	STEP 3. The severity index of the stay is potentially increased by 1 depending on the root-DRG and the age of the patient
PD	N39.0	N39.0	N39.0
Root DRG	11M04	11M04	11M04
Age	82	82	82
LOS	4	4	4
Secondary diagnoses (SD)	~~B96.2~> C78.0 Z29.0 U82.2 A41.5	~~B96.2~> C78.0 1 Z29.0 3 U82.2 3 A41.5 4	~~B96.2~> C78.0 1 Z29.0 3 U82.2 3 A41.5 4

STEP 4. The severity index of the stay is potentially decreased if the LOS is lower than a specified threshold:
3 days min. for lvl 2
4 days min. for lvl 3
5 days min. for lvl 4

→ Level 3

Ways of improvement
Inspired by others…

FROM…

STEP 2. The severity index of the stay is the max level of its SD

N39.0

~~B96.2~~	
C78.0	1
Z29.0	3
U82.2	3
A41.5	4

TO…

Classical SD model

STEP 2. The severity index of the stay is the max level of its SD…but we increase the number of levels from 4 to 5

~~B96.2~~	
C78.0	1
Z29.0	4
U82.2	4
A41.5	5

Max= 5

PD-conditionned SD model

STEP 2. The severity index of the stay is the max level of its secondary diagnoses…*but the level of each SD is dependent on the PD*

~~B96.2~~	
C78.0 (with PD=N39.0)	1
Z29.0 (with PD=N39.0)	2
U82.2 (with PD=N39.0)	4
A41.5 (with PD=N39.0)	3

Max= 4

Multiple SD model

STEP 2. The severity index of the stay ~~is the max level of its secondary diagnoses~~ *is determined by the number of secondary diagnoses per level*

~~B96.2~~	
C78.0	1
Z29.0	3
U82.2	3
A41.5	4

Matrix Combination
(one final level per combination)

#1-0-2-1 → 4

Lvl 1 SD : 1
Lvl 2 SD : 0
Lvl 3 SD : 2
Lvl 4 SD : 1

Method
Optimization process

- Each diagnosis and combination rule is given a level={1,…N} such as to maximize R^2(LOS,DRG)
- Because ALOS of a DRG is a discrete function, it is impossible to use derivative-based optimization procedures or classical statistical tools to determine optimal levels
- Optimization is thus conducted using a hill-climbing method developped by ATIH
  - The method has for goal to give the **optimal** level to each SD (as well as combination rule for mutliple-SD model) w.r.t R^2
- Results are then checked by physicians to ensure consistency

Optimization
Maximum SD

AGENCE TECHNIQUE DE L'INFORMATION SUR L'HOSPITALISATION

Current levels

R^2 is computed for each admissible severity level of the current code

The level maximizing the R^2 is memorized

The two previous steps are repeated for each code

Once all codes have been parsed, the modification yielding the strongest improvement is kept

			Best	ΔR^2
A00.0	2	~~2~~ 1,3,4	1	+0,00011
A00.1	2			
A00.9	2			
…		…		
Z99.8	1	Z99.8 1		
Z99.9	1	Z99.9 1		

			Best	ΔR^2
A00.0	2		1	+0,00021
A00.1	~~2~~ 1,3,4		1	+0,00001
A00.9	2			
…		…		
Z99.8	1	Z99.8 1		
Z99.9	1	Z99.9 1		

A00.0	1	1	+0,00011
A00.1	2		
A00.9	2		
…			
Z99.8	1		
Z99.9	1		

Updated levels

Re-iterate until no single modification improves the R^2

Optimization
Other models

- The same optimization process is used for all types of models
- The changes lie in the way inputs (diags and combination rules) are presented to the algorithm

PD-conditionned model

PD	SD	
A00.1	A00.0	2
A00.9	A00.0	2
A01.0	A00.0	2
…		
Z99.8	Z99.9	2

A level is given to each combination of PD-SD

PD is replaced by OTHER for rare combinations (<100 stays per year)

Multiple-SD model

Combination rules	#1-0-0-0	1
	#2-0-0-0	1
	#3-0-0-0	1
	#4-0-0-0	1
	#0-1-0-0	2
	#0-2-0-0	3
	…	
	#4-4-4-4	4
Diagnoses	A00.0	2
	A00.1	2
	A00.9	2
	…	
	Z99.8	1
	Z99.9	1

The levels of combination rules and SD are « jointly » optimized

In a stay, related SD must be taken into account. We proceed as follows :

1- Sort SD by decreasing levels
2- Any SD excluded by a (non previously excluded) higher ranked SD is excluded form the stay

This process is repeated at each iteration because of the potential changes in SD levels

Tested models and preliminary results

- Optimization (i.e, determination of optimum severity levels of diagnoses and their combinations for mutliple-SD) has been conducted using the 2017 + 2018 French National Databases
- Results (presented hereafter) are computed on the 2019 French National Database

Model	# SD levels	# severity index levels	R^2 (LOS,GHM)	Difference with baseline
Maximum SD	*4*	*4*	*41.7%*	*0%*

Tested models and preliminary results

- Optimization (i.e, determination of optimum severity levels of diagnoses and their combinations for mutliple-SD) has been conducted using the 2017 + 2018 French National Databases
- Results (presented hereafter) are computed on the 2019 French National Database

Model	# SD levels	# severity index levels	R^2 (LOS,GHM)	Difference with baseline
Maximum SD	*4*	*4*	*41.7%*	*0%*
Maximum SD	5	5	42.6%	+0.9%

Tested models and preliminary results

- Optimization (i.e, determination of optimum severity levels of diagnoses and their combinations for mutliple-SD) has been conducted using the 2017 + 2018 French National Databases
- Results (presented hereafter) are computed on the 2019 French National Database

Model	# SD levels	# severity index levels	R^2 (LOS,GHM)	Difference with baseline
Maximum SD	*4*	*4*	*41.7%*	*0%*
Maximum SD	5	5	42.6%	+0.9%
Maximum SD (PD-c)	4	4	42.7%	+1.0%

Tested models and preliminary results

- Optimization (i.e, determination of optimum severity levels of diagnoses and their combinations for mutliple-SD) has been conducted using the 2017 + 2018 French National Databases
- Results (presented hereafter) are computed on the 2019 French National Database

Model	# SD levels	# severity index levels	R^2 (LOS,GHM)	Difference with baseline
Maximum SD	*4*	*4*	*41.7%*	*0%*
Maximum SD	5	5	42.6%	+0.9%
Maximum SD (PD-c)	4	4	42.7%	+1.0%
Multiple SD	4	4	46.3%	+4.6%

Tested models and preliminary results

- Optimization (i.e, determination of optimum severity levels of diagnoses and their combinations for mutliple-SD) has been conducted using the 2017 + 2018 French National Databases
- Results (presented hereafter) are computed on the 2019 French National Database

Model	# SD levels	# severity index levels	R^2 (LOS,GHM)	Difference with baseline
Maximum SD	*4*	*4*	*41.7%*	*0%*
Maximum SD	5	5	42.6%	+0.9%
Maximum SD (PD-c)	4	4	42.7%	+1.0%
Multiple SD	4	4	46.3%	+4.6%
Multiple SD	4	5	48.0%	+6.3%

Tested models and preliminary results

- Optimization (i.e, determination of optimum severity levels of diagnoses and their combinations for mutliple-SD) has been conducted using the 2017 + 2018 French National Databases
- Results (presented hereafter) are computed on the 2019 French National Database

Model	# SD levels	# severity index levels	R^2 (LOS,GHM)	Difference with baseline
Maximum SD	*4*	*4*	*41.7%*	*0%*
Maximum SD	5	5	42.6%	+0.9%
Maximum SD (PD-c)	4	4	42.7%	+1.0%
Multiple SD	4	4	46.3%	+4.6%
Multiple SD	4	5	48.0%	+6.3%

Tested models and preliminary results

Model	#SD levels	#severity index levels	% lvl 1	% lvl 2	% lvl 3	% lvl 4	% lvl 5	ALOS lvl1	ALOS lvl2	ALOS lvl3	ALOS lvl4	ALOS lvl5
Maximum SD	4	4	54.3	25.9	15.9	3.9	-	2.8	6.7	12.1	21.2	-
Multiple SD	4	4	57.2	25.7	15.5	1.4	-	2.9	6.8	13.9	30.4	-

Tested models and preliminary results

Model	#SD levels	#severity index levels	% lvl 1	% lvl 2	% lvl 3	% lvl 4	% lvl 5	ALOS lvl1	ALOS lvl2	ALOS lvl3	ALOS lvl4	ALOS lvl5
Maximum SD	4	4	54.3	25.9	15.9	3.9	-	2.8	6.7	12.1	21.2	-
Multiple SD	4	4	57.2	25.7	15.5	1.4	-	2.9	6.8	13.9	30.4	-
Maximum SD	5	5	53.4	24.2	13.9	6.2	2.3	2.7	6.3	10.8	15.6	23.5
Multiple SD	4	5	51.8	26.3	15.6	5.5	0.6	2.7	6.0	11.3	18.9	37.9

Pros and cons of each model

MODEL	PROS	CONS
Maximum SD	- Well-known by stakeholders	No more potential in improving the quality of DRGs, except increasing the nb of severity levels
Maximum SD (PD-conditionned)	- Room for fairly good improvement of DRGs - Appreciated theoretically by physicians	When looking in details, some results are not easily understandable (hard to understand/justify why some SD have different levels across specific PD) Requires a lot of regularization (manual parsing of results) to ensure that: - close SD have similar levels given a PD - Close PD give similar levels to similar SD
Multiple SD	- Asked for a long time by hospitals representatives - Very intuitive : the more severe diseases there are, the more severe the stay - Room for strong improvement of DRGs	- Optimization is technically hard - Necessity to deal with SD exclusions - Suspicions of SD under-coding in national database => biased estimations ?

Conclusion and future work

- Changing the way of defining severity index seems promising in terms of classification performance
  - The first results suggest that Multiple-SD >> PD-conditonned ≈ 5 levels
- Searching for the optimal results by a dedicated algorithm gives an upper bound on the R^2 attainable by each model
- Following steps
  - Physicians must « regularize » the different models, i.e, analyzing the results given by the optimization procedure
  - Translating brute optimization results to understandable and readable results (for physicians) is a real challenge
    - Tools for listing and stays' analysis are under development
  - Impact of each model must be precised in order for stakeholders to clearly understand the consequences of the changes

atih
AGENCE TECHNIQUE
DE L'INFORMATION
SUR L'HOSPITALISATION

Thanks for your attention

Agence technique de l'information sur l'hospitalisation

117, bd Marius Vivier Merle 69329 Lyon cedex 03
Téléphone : 04 37 91 33 10
Fax : 04 37 91 33 67

www.atih.sante.fr

Appendix

Method
Tips for optimization

- 3 types of tips and tricks
  - Structure data in order to optimize speed of computations
    - Dependent on the infrastructure and softwares
  - Reduce the dimensionality of the variables using medical knowledge
    - Gathering ICD-10 codes
      - Some ICD-10 codes are required to change together, for example all codes of the category A00 are required to have the same level
      - In such cases all codes starting with A00 may be gathered into only one group “A00”
      - The 17,000 ICD-codes are gathered into 3,000 groups
    - Define a set of admissible levels for each ICD-10 codes
      - Imprecise codes for examples are usually not allowed to have a severity level>2
      - Saves some computations
  - Algorithmic approximations
    - Start the process with the best possible approximation of severity levels (good initialization)
    - At each iteration, test only adjacent levels of the current one (max. 2 tests/code)
    - Allow sometimes more than only one modification at each iteration, especially during the first ones
      - A parameter evolving at each iteration gives an upper threshold of the number of simultaneous modifications allowed

Multiple SD in acute care
4 levels of diag / 4 levels of severity index

AGENCE TECHNIQUE DE L'INFORMATION SUR L'HOSPITALISATION

Rules obtained with hill-climbing otpimization
(Diagnoses of severity level 1 are never used to modulate severity)

severity level of the stay		Nb of diag of level 2				
Nb of diag of level 4	Nb of diag. Of level 3	0	1	2	3	4 or more
0	0	1	2	2	2	3
	1	2	2	3	3	3
	2	3	3	3	3	4
	3	3	3	4	4	4
	4	4	4	4	4	4
1	0	3	3	3	3	4
	1	3	3	4	4	4
	2	4	4	4	4	4
	3	4	4	4	4	4
	4	4	4	4	4	4
2 or more	0	4	4	4	4	4
	1	4	4	4	4	4
	2	4	4	4	4	4
	3	4	4	4	4	4
	4	4	4	4	4	4

residuals		Nb of diag of level 2				
Nb of diag of level 4	Nb of diag of level 3	0	1	2	3	4 or more
0	0	- 1,7	- 0,7	0,7	2,2	4,2
	1	0,3	2,3	3,9	5,3	7,6
	2	4,4	6,1	7,7	9,1	12,1
	3	9,2	10,9	12,4	14,4	18,0
	4	16,7	18,3	20,5	22,7	28,3
1	0	6,0	5,0	7,1	9,2	13,3
	1	8,4	10,4	12,7	14,9	19,7
	2	14,7	16,8	19,0	21,6	27,7
	3	20,6	24,0	26,4	29,4	37,2
	4	33,4	36,0	40,6	41,1	52,0
2 or more	0	14,7	13,8	17,0	20,1	27,9
	1	18,3	20,5	23,5	28,1	33,9
	2	26,3	28,1	34,5	34,2	43,3
	3	39,1	41,7	42,3	41,6	55,5
	4	48,5	62,8	60,8	64,3	72,6

Residuals = LOS – ALOS of the root DRG

Multiple SD in acute care
4 levels of diag / 5 levels of severity index

atih
AGENCE TECHNIQUE DE L'INFORMATION SUR L'HOSPITALISATION

Rules obtained with hill-climbing otpimization
(Diagnoses of severity level 1 are never used to modulate severity

Nb of diag of level 4	Nb of diag of level 3	Nb of diag of level 2				
		0	1	2	3	4 or more
0	0	1	2	2	3	3
	1	2	3	3	3	3
	2	3	3	3	4	4
	3	3	4	4	4	4
	4	4	4	4	4	4
1	0	3	3	3	4	4
	1	3	4	4	4	4
	2	4	4	4	4	4
	3	4	4	4	4	5
	4	4	5	5	5	5
2	0	4	4	4	4	4
	1	4	4	4	4	5
	2	4	4	5	5	5
	3	5	5	5	5	5
	4	5	5	5	5	5
3	0	4	5	5	5	5
	1	5	5	5	5	5
	2	5	5	5	5	5
	3	5	5	5	5	5
	4	5	5	5	5	5
4	0	5	5	5	5	5
	1	5	5	5	5	5
	2	5	5	5	5	5
	3	5	5	5	5	5
	4	5	5	5	5	5

Nb of diag of level 4	Nb of diag of level 3	Nb of diag of level 2				
		0	1	2	3	4 or more
0	0	-1,7	-1,0	0,1	1,3	2,8
	1	-0,3	1,2	2,5	3,6	5,0
	2	2,1	3,6	4,8	5,9	7,4
	3	4,9	6,3	7,4	8,5	10,2
	4	8,8	10,2	11,6	12,7	15,2
1	0	2,8	3,4	5,0	6,4	8,1
	1	4,8	6,7	8,0	9,1	10,8
	2	8,4	9,9	11,2	12,2	14,2
	3	12,3	13,4	14,7	16,3	18,9
	4	17,9	20,0	22,8	24,5	27,9
2	0	9,1	10,7	12,0	13,2	15,0
	1	13,1	14,5	15,5	16,6	19,4
	2	17,1	18,3	20,2	21,3	24,1
	3	22,9	24,5	25,6	26,7	30,8
	4	30,2	32,3	33,9	37,6	42,7
3	0	17,3	20,2	20,8	20,4	24,5
	1	22,7	24,5	26,2	27,7	28,7
	2	27,5	29,6	30,4	33,1	36,3
	3	35,7	35,7	40,0	40,7	43,9
	4	46,7	48,4	51,4	51,0	57,0
4	0	27,0	29,7	44,0	50,0	32,5
	1	32,5	37,9	43,4	43,5	43,5
	2	43,1	47,1	53,3	59,8	50,9
	3	46,6	52,4	55,1	51,0	66,2
	4	63,0	72,5	66,3	72,4	80,7