

Performance of Mortality prognostic scores in Portuguese PICU's

Introdução à Medicina II

CLASS 4¹; OLIVEIRA, R. C. S.²

¹ turma4fmup09@gmail.com; ² rcoliveira@med.up.pt (Adviser)

Performance of Mortality prognostic scores in Portuguese PICU's

“Are they doing a good job?”

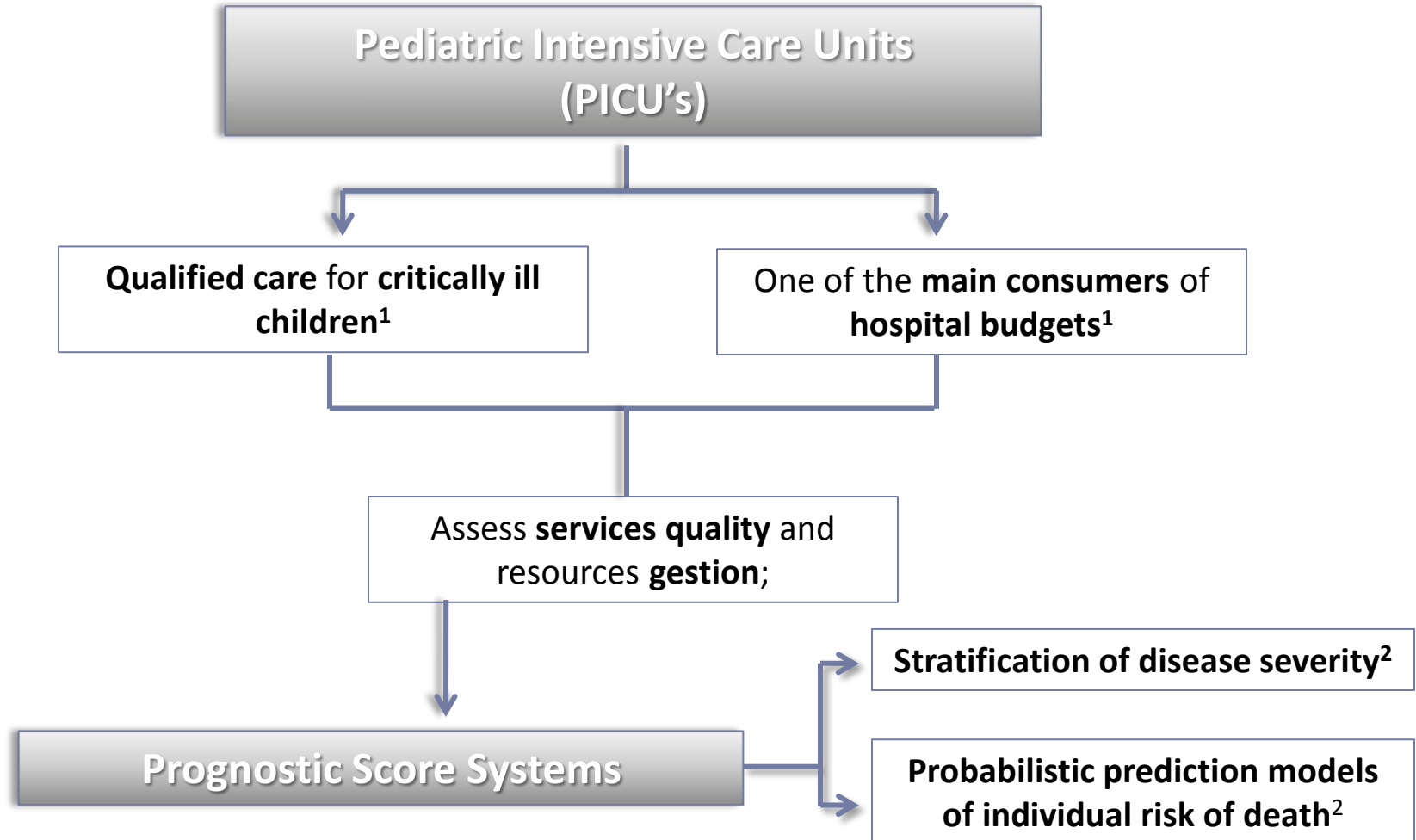
TABLE OF CONTENTS

- ▶ **Introduction**
 - ▶ Context
 - ▶ General Aims
- ▶ **Methods**
 - ▶ Study Design
 - ▶ Data Acquisition
 - ▶ Algorithms calculation
 - ▶ Statistical analysis
 - ▶ Discrimination
 - ▶ Calibration
- ▶ **Results**
 - ▶ Preliminary discussion
- ▶ **Project amplification**

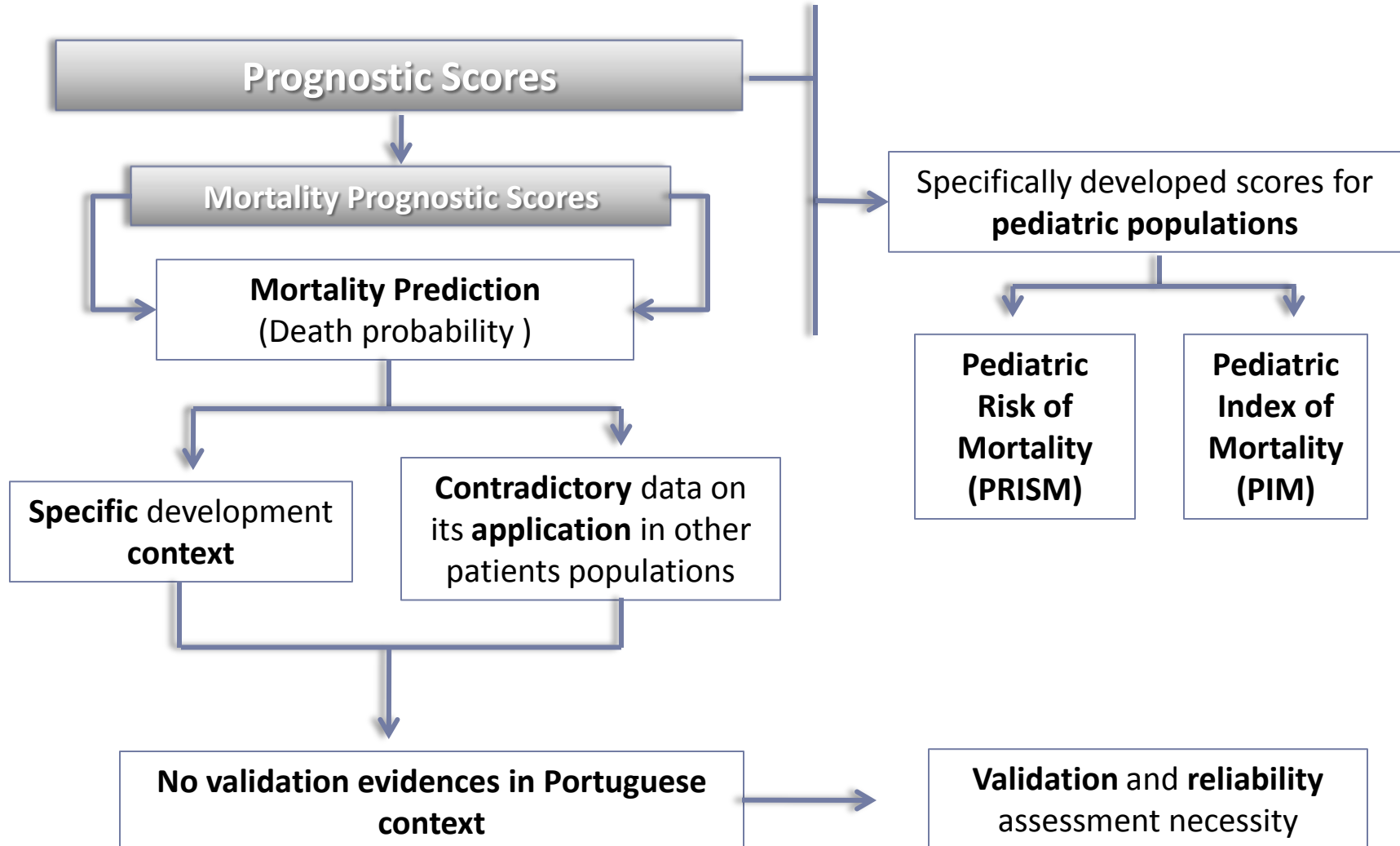
TABLE OF CONTENTS

- ▶ **Introduction**
 - ▶ Context
 - ▶ General Aims
- ▶ **Methods**
 - ▶ Study Design
 - ▶ Data Aquisition
 - ▶ Algorithms calculation
 - ▶ Statistical analysis
 - ▶ Discrimination
 - ▶ Calibration
- ▶ **Results**
 - ▶ Preliminary discussion
- ▶ **Project amplification**

INTRODUCTION



INTRODUCTION



Assessment and Optimization of Mortality Prediction Tools for Admissions to Pediatric Intensive Care in the United Kingdom

ABSTRACT

OBJECTIVE. To assess the Pediatric Risk of Mortality (PRISM, PRISM III-12, and PRISM III-24) systems and the Pediatric Index of Mortality (PIM and PIM2) systems for use in comparing the risk-adjusted mortality of children after admission for pediatric intensive care in the United Kingdom.

METHODS. All PICUs in the United Kingdom were invited to participate. Predicted probability of PICU mortality was calculated using the published algorithms for PIM, PIM2, and PRISM and compared with observed mortality. These scores, along with PRISM III-12 and PRISM III-24, whose algorithms are not published, were optimized for the United Kingdom.

RESULTS. Of 26 PICUs in the United Kingdom, 22 (85%) were recruited, and sufficient prospective data were collected from 18 (69%) units on 10 197 (98%) of 10 385 admissions between March 2001 and February 2002. All published tools were found to have poor calibration but provided good discriminatory power. After estimation of UK-specific coefficients, only PIM2, PRISM III-12, and PRISM III-24 had satisfactory calibration. All models provided good discriminatory power. Funnel plots for all of the recalibrated models indicated that the risk-adjusted mortality for all units was consistent with random variation.

CONCLUSIONS. PIM2, PRISM III-12, and PRISM III-24 all were found to be suitable for use in a UK PICU setting. All tools provided similar conclusions in assessing the distribution of risk-adjusted mortality in UK PICUs. It now is important that these tools be used to monitor outcome and improve the quality of pediatric intensive care within the United Kingdom.

Pediatric Index of Mortality 2 score in Italy: a multicenter, prospective, observational study

Abstract *Objectives:* To assess the performance of the Pediatric Index of Mortality (PIM) 2 score in Italian pediatric intensive care units (PICUs). *Design:* Prospective, observational, multicenter, 1-year study. *Setting:* Eighteen medical–surgical PICUs. *Patients:* Consecutive patients (3266) aged 0–16 years admitted between 1 March 2004 and 28 February 2005. *Interventions:* None. *Measurements and main results:* To assess the performance of the PIM2 score, discrimination and calibration measures were applied to all children admitted to the 18 PICUs, in the entire population and in different groups divided for deciles of risk, age and admission diagnosis. There was good discrimination, with an area under

the receiver operating characteristic (ROC) curve of 0.89 (95% CI 0.86–0.91) and good calibration of the scoring system [non-significant differences between observed and predicted deaths when the population was stratified according to deciles of risk (χ^2 9.86; 8 df, $p = 0.26$) for the whole population]. *Conclusions:* The PIM2 score performed well in this sample of the Italian pediatric intensive care population. It may need to be reassessed in the injury and postoperative groups in larger studies.

Keywords Pediatric intensive care unit · Severity score · Mortality · Pediatric index of Mortality · Children

- ▶ **Assessment to Pediatric Risk of Mortality (PRISM, PRISM III) and Pediatric Index of Mortality (PIM and PIM2) systems for use in comparing the risk-adjusted mortality of children after admission for pediatric intensive care in Portugal;**
- ▶ Validation of PRISM, PRISM III, PIM and PIM-2 prognostic scores.
 - ▶ Comparing their **performance** at a general Portuguese Pediatric Intensive Care Units;
- ▶ Statistical evaluation of PRISM, PRISM III, PIM and PIM-2 scoring systems' discrimination, calibration and predictive degree at Portuguese PICU's.

- ▶ Assessment to **Pediatric Risk of Mortality** (PRISM, PRISM III) and **Pediatric Index of Mortality** (PIM and PIM2) systems for use in **comparing the risk-adjusted mortality** of children after admission for pediatric intensive care in Portugal;
- ▶ **Validation** of **PRISM**, **PRISM III**, **PIM** and **PIM-2** prognostic scores.
 - ▶ Comparison of their **performance** at a general Portuguese Pediatric Intensive Care Units;
- ▶ Statistical evaluation of **PRISM**, **PRISM III**, **PIM** and **PIM-2** scoring systems **discrimination, calibration and predictive degree** at Portuguese PICU's.

- ▶ Assessment to **Pediatric Risk of Mortality** (PRISM, PRISM III) and **Pediatric Index of Mortality** (PIM and PIM2) systems for use in **comparing the risk-adjusted mortality** of children after admission for pediatric intensive care in Portugal;
- ▶ Validation of PRISM, PRISM III, PIM and PIM-2 prognostic scores.
 - ▶ Comparison of their **performance** at a general Portuguese Pediatric Intensive Care Units;
- ▶ **Statistical evaluation** of PRISM, PRISM III, PIM and PIM-2 scoring systems' **discrimination, calibration and predictive degree** at Portuguese PICU's;

TABLE OF CONTENTS

- ▶ Introduction
 - ▶ Context
 - ▶ General Aims
- ▶ **Methods**
 - ▶ Study Design
 - ▶ Data Acquisition
 - ▶ Algorithms calculation
 - ▶ Statistical analysis
 - ▶ Discrimination
 - ▶ Calibration
- ▶ Results
 - ▶ Preliminary discussion
- ▶ Project amplification

Validation

- ▶ An integrated evaluative judgment of the degree to which **empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences** and actions based on test scores or other modes of assessment.⁴

I) Data acquisition

II) Algorithms calculation

III) Statistical validity assessment

→ Via SPSS and other Microsoft Office resources

Standard Criteria

- Discrimination
- Calibration
- Explanatory power

Bases of defined strategy

- ▶ Precursor project – **REUNIR** (*Recolha Uniformizada e Nacional de Informação Relevante*);
- ▶ Conducted Study in UK (same thematic and analytical bases);⁵
- ▶ Other Prognostic Scores Validation Studies in Portugal [e.g. **APPACHE** (*Acute Physiology, Age, Chronic Health Evaluation*), **SABS** (*Clinical Risk Index for Babies*)];⁶

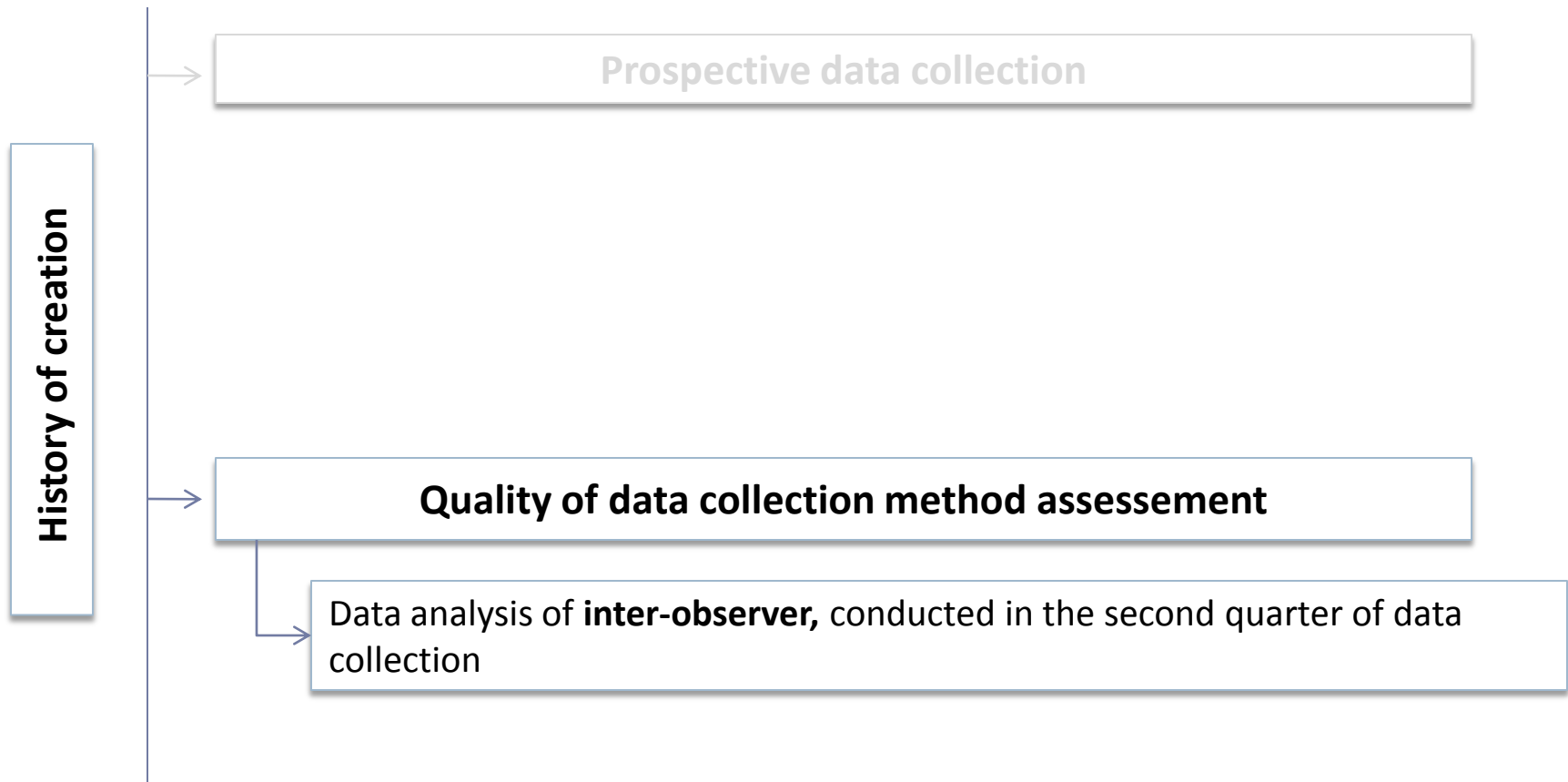
I) Previously created database

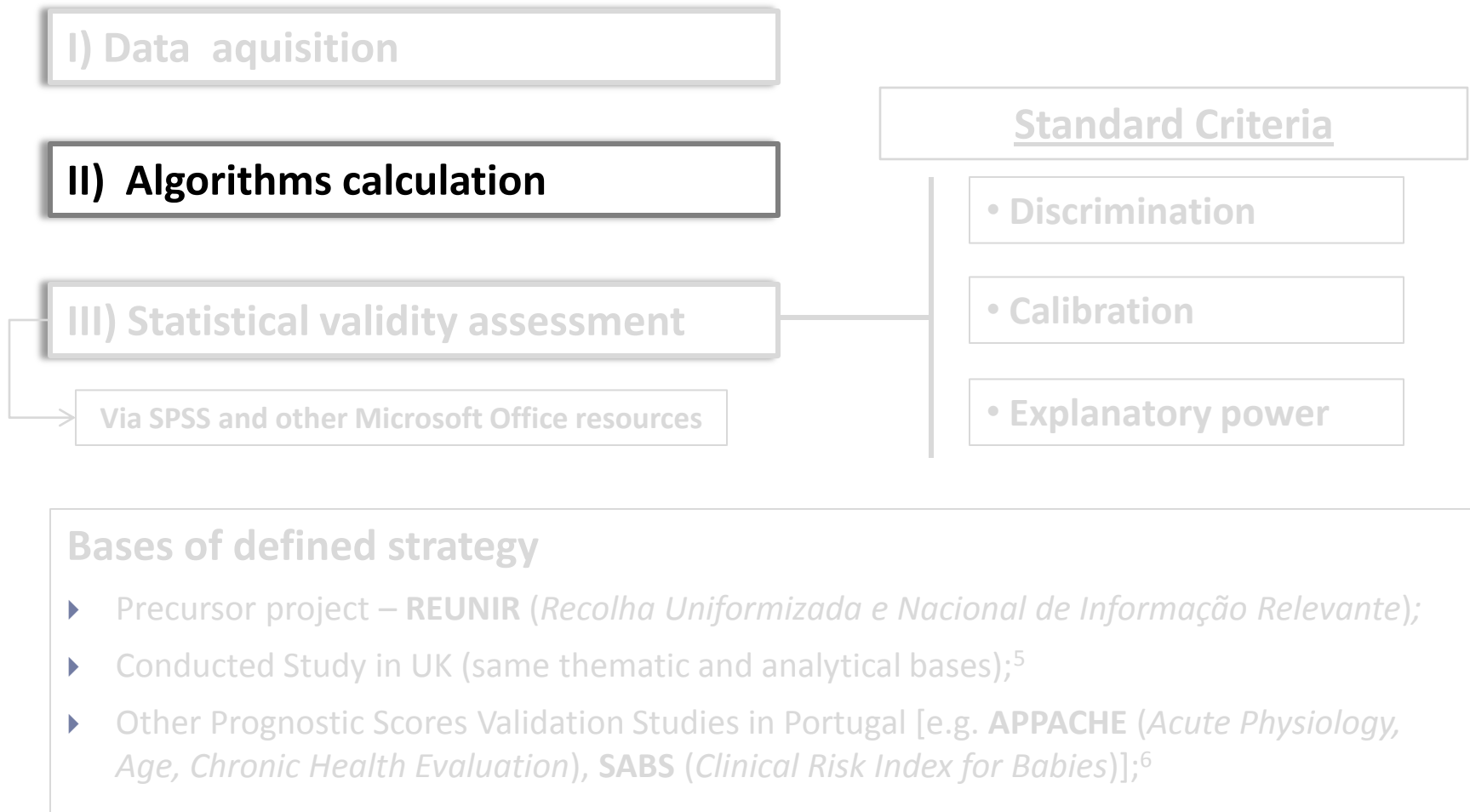
History of creation

Prospective data collection

- **Time of collection:** 30 months
- **Institutions of collection:** 3 volunteers Portuguese PICU's (Hospital Pediátrico de Coimbra - Coimbra, Hospital D. Estefânia - Lisbon, Hospital São João - Oporto)
- **Dimension:** 2000 patients
- **Inclusion / Exclusion criteria:** All admissions between 29 days and 16 years old; No more criteria are known;
- **Data:** All necessary data for PIM, PIM II, PRISM, PRISM III calculation (Routinely collection; Added pro-form);

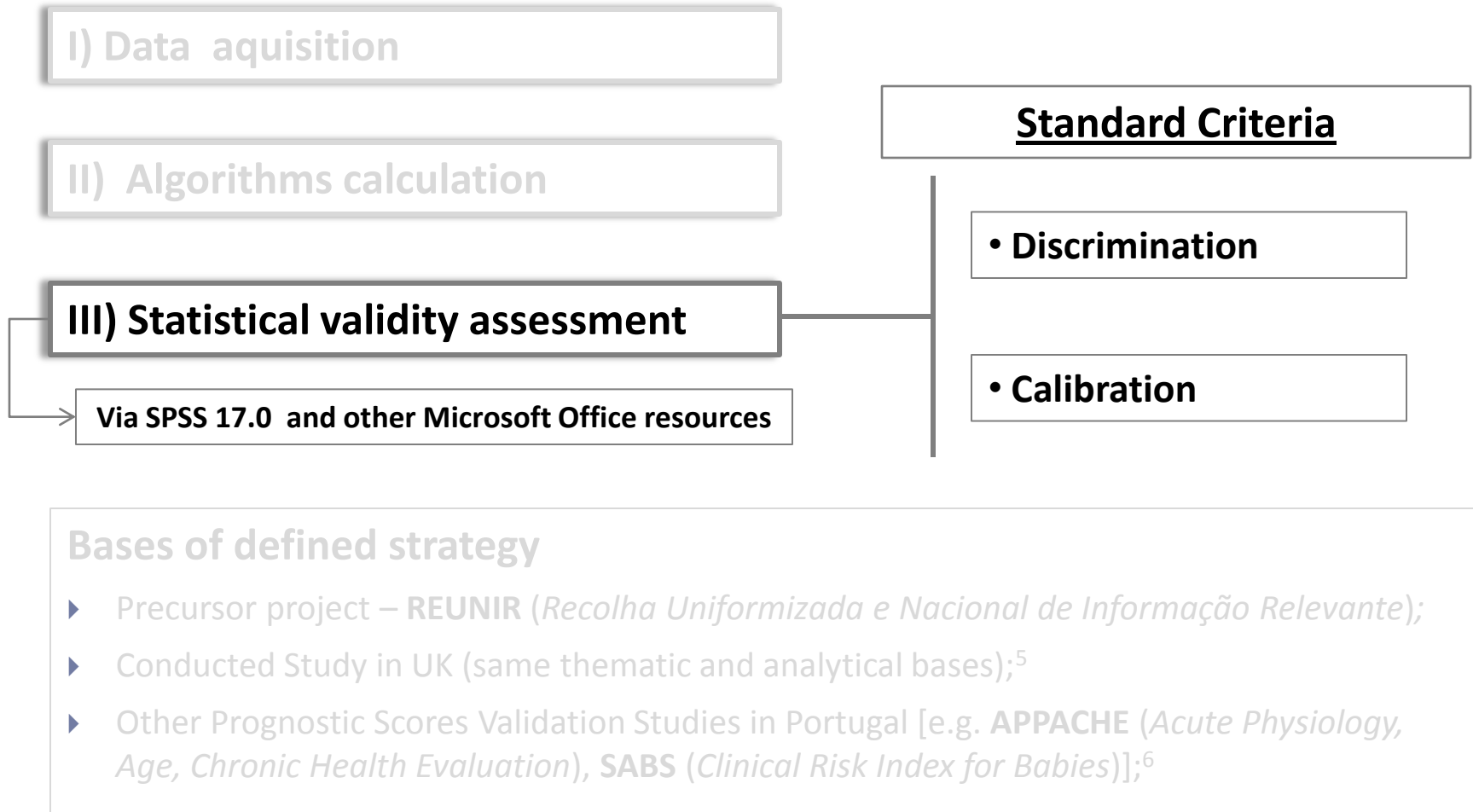
I) Previously created database



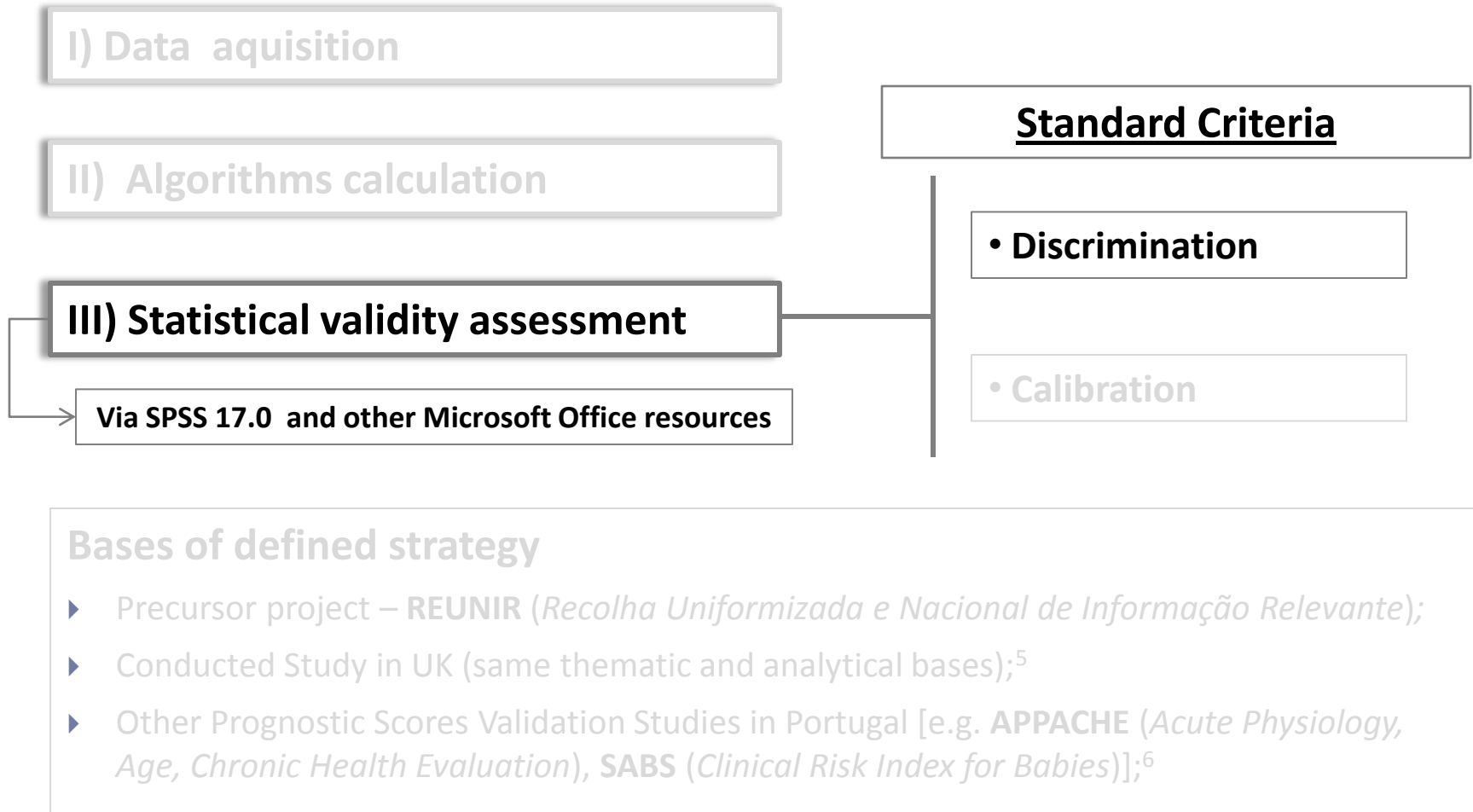


ALGORITHMS CALCULATION

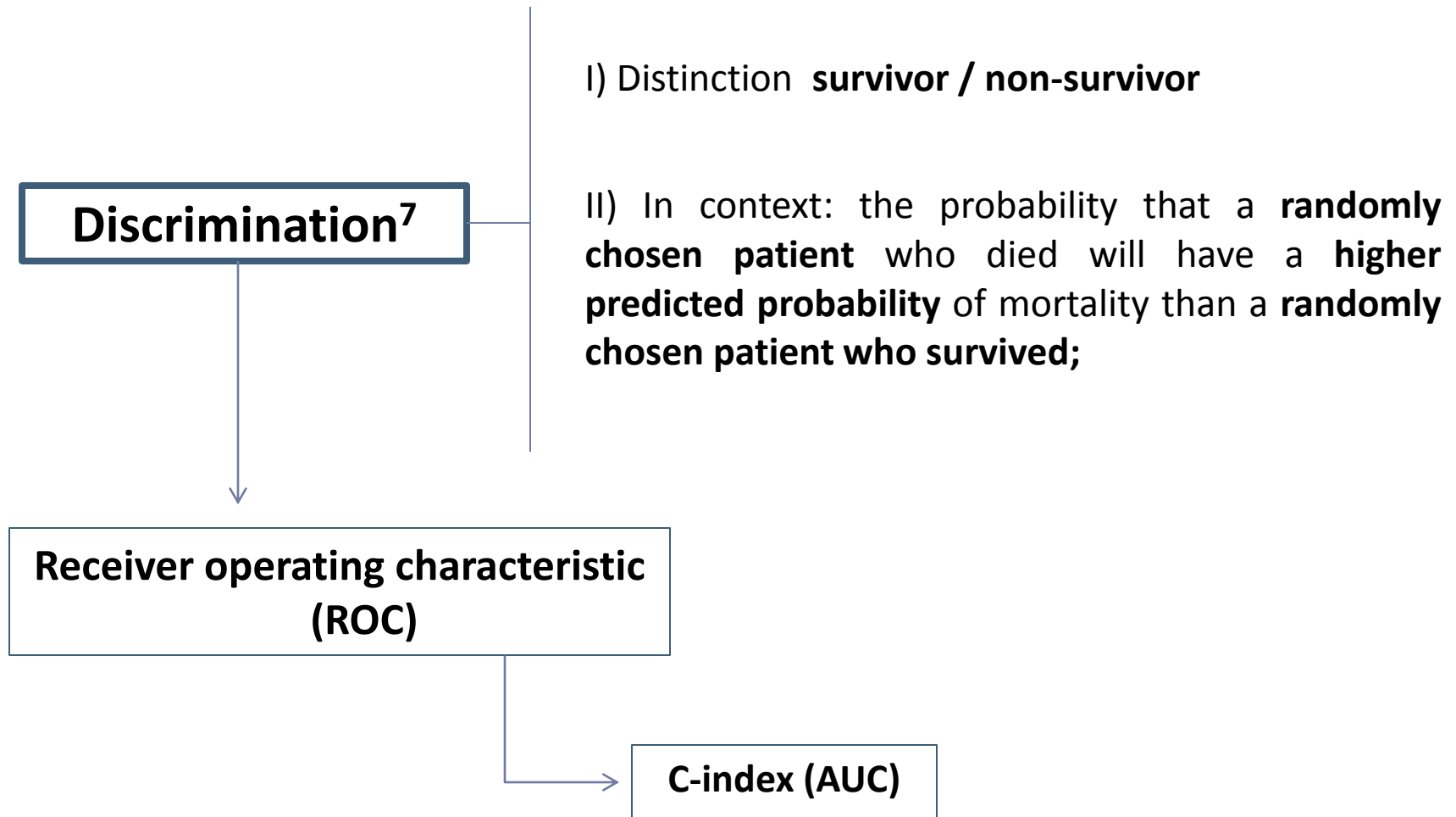
- I) According to **published equations**;
 - II) Informatic **software applications** of Pediatric Mortality Prognostic Scores calculation;
-



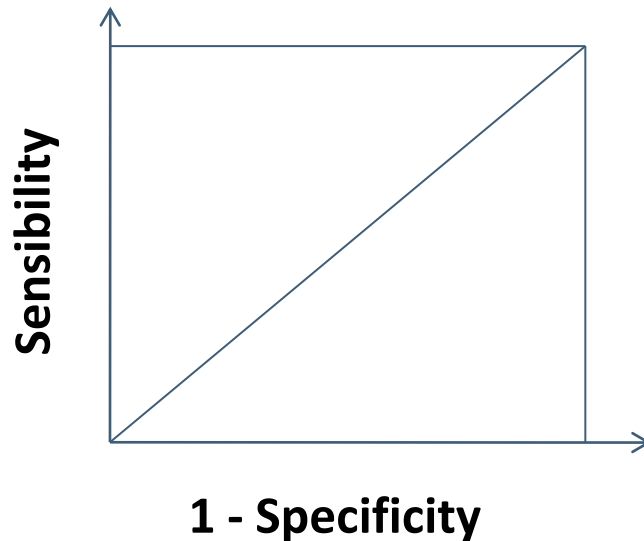
STUDY DESIGN



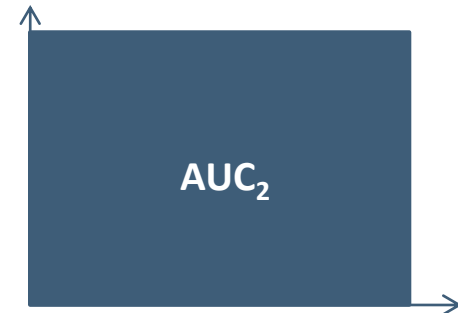
DISCRIMINATION



ROC Curves



Lowest discrimination power

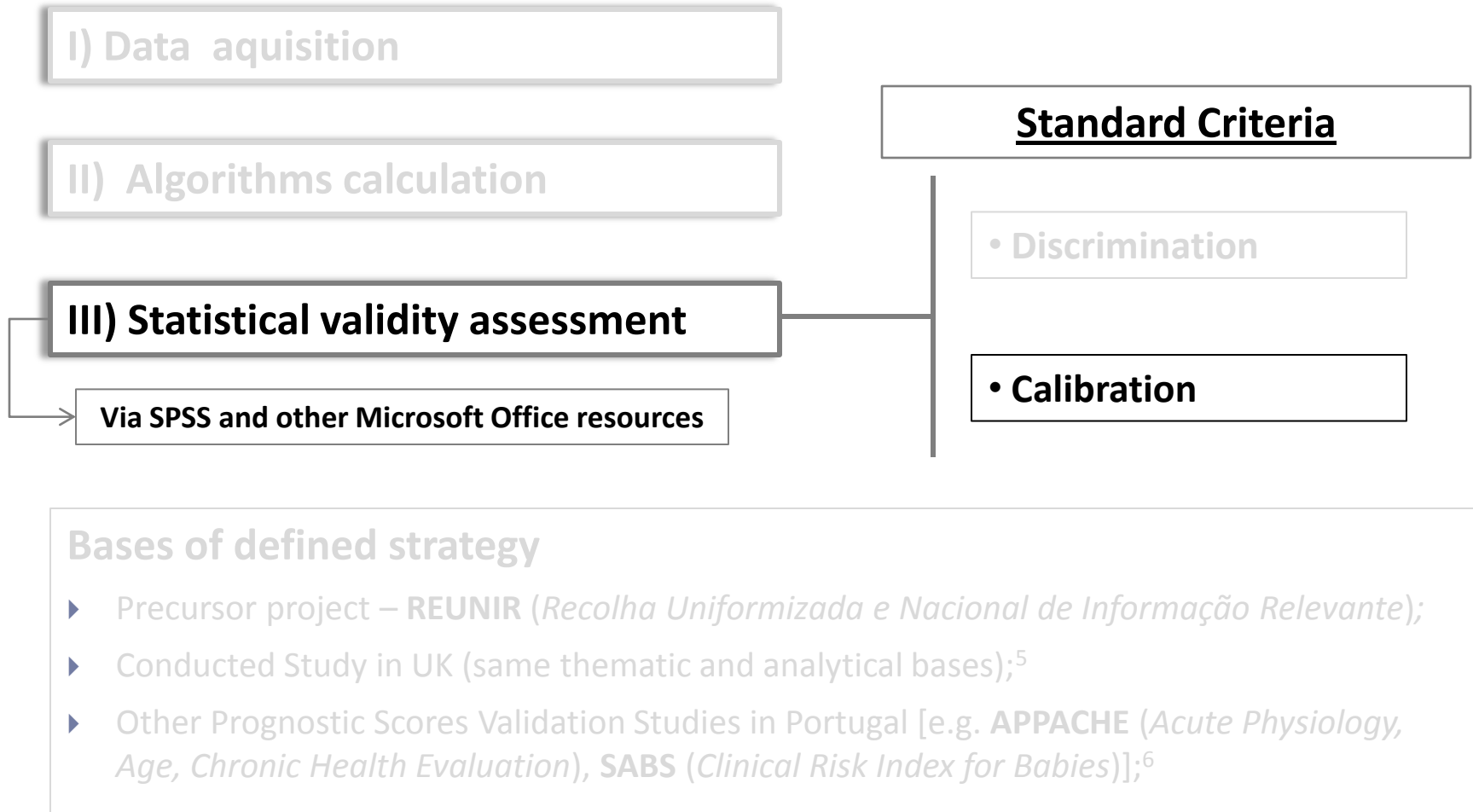


Highest discrimination power

C-index (AUC)²

- ≥ 0.7 acceptable
- ≥ 0.8 good
- ≥ 0.9 excellent

STUDY DESIGN



Calibration⁸

The ability of a model to **match predictive and observed** death rates across the entire spread of data;

Calibration⁸



Goodness-of-fit Hosmer
Lemeshow test

- I. **Classification** into $g=10$ (or possibly less) decile of **risk groups** based on the values of the **estimated probabilities**;
- II. **Common X^2 test** for the mean of the **predicted probability** against the **observed fraction of events**;
- III. **Null hypothesis:** No differences in observed and expected number of deaths;

TABLE OF CONTENTS

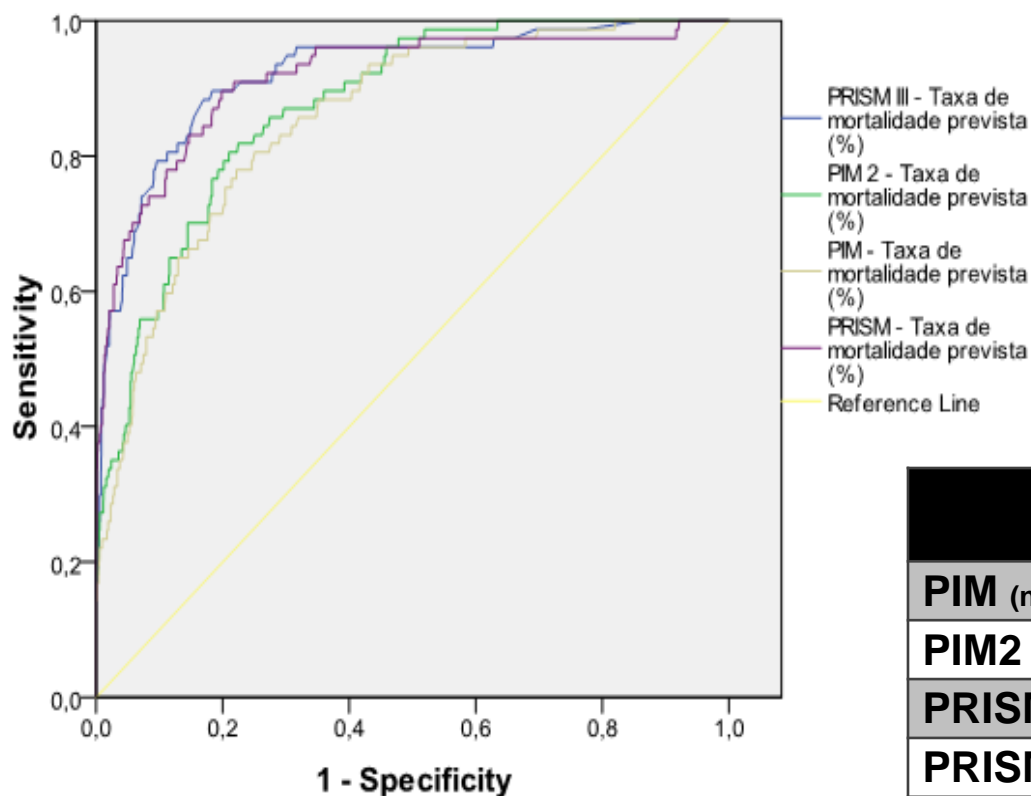
- ▶ **Introduction**
 - ▶ Context
 - ▶ General Aims
- ▶ **Methods**
 - ▶ Study Design
 - ▶ Data Aquisition
 - ▶ Algorithms calculation
 - ▶ Statistical analysis
 - ▶ Discrimination
 - ▶ Calibration
- ▶ **Results**
 - ▶ Preliminary discussion
- ▶ **Project amplification**

General characteristics of data sample

Characteristics	Value
Number of patients	1809
Mortality	155 (8,6%)
Age (media)	4,6 years
Internament duration	7,7 days
Gender: male	54%
Admission reason	
• Surgery	37,3%
• Medical	59,9%
• Monitorization and prevention	2,9%
Mechanical ventilation during internament	57,4%

Discrimination assessment: Receiver Operating Characteristics

ROC Curve



Score	AUC	CI 95%
PIM (n = 1809)	0,84	[0,81;0,87]
PIM2 (n = 1809)	0,89	[0,85;0,92]
PRISM (n = 1809)	0,90	[0,87;0,92]
PRISM III (n = 1809)	0,91	[0,88;0,93]

PRISM - Pediatric Risk of Mortality; PIM – Pediatric Index of Mortality; AUC – Area under curve

Calibration assessment: Hosmer-Lemeshow goodness-of-fit test

PRISM III – An example

Contingency Table for Hosmer and Lemeshow Test

	Estado clínico na alta da UCIP = Vivo		Estado clínico na alta da UCIP = Falecido		Total
	Observed	Expected	Observed	Expected	
Step 1 1	900	897,499	6	8,501	906
2	455	459,121	19	14,879	474
3	173	172,836	19	19,164	192
4	124	122,545	97	98,455	221

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	1,961	2	,375

RESULTS

Summarizing:

Score	Area Under ROC Curve (AUC)	Chi-Square (8df)	p
PIM (n = 1809)	0,84	4,05	0,132
PIM2 (n = 1809)	0,89	7,23	0,027
PRISM (n = 1809)	0,90	3,62	0,305
PRISM III (n = 1809)	0,91	1,96	0,375

PRISM - Pediatric Risk of Mortality; PIM – Pediatric Index of Mortality;

RESULTS

Analysis by PICU at a glance

PICU	Score	Observed mortality rate	Expected mortality rate	SMR
A	PIM	8,2	6,7	1,22
	PIM-2	8,2	5,7	1,44
	PRISM	8,2	10,2	0,80
	PRISM III	8,2	7,2	1,14
B	PIM	5,5	4,8	1,15
	PIM-2	5,5	3,8	1,45
	PRISM	5,5	5,3	1,04
	PRISM III	5,5	4,3	1,28
C	PIM	12,4	8,8	1,41
	PIM-2	12,4	8,2	1,51
	PRISM	12,4	15,3	0,81
	PRISM III	12,4	11,5	1,08

PRISM - Pediatric Risk of Mortality; PIM – Pediatric Index of Mortality; PICU – Pediatric Intensive Care Unit ; SMR – Standardized mortality rate

Analysis by PICU at a glance

Comparative Audit purpose

SMR – Standardized Mortality Rate;

- Obtained **by comparing the observed mortality** in the population **with the expected mortality** which have occurred had the standard rates applied;

It is assumed that:

- **SMR > 1.0** may reflect poor care;
- **SMR <1.0** may reflect good care;

SMR
1,22
1,44
0,80
1,14
1,15
1,45
1,04
1,28
1,41
1,51
0,81
1,08

Analysis by PICU at a glance

PICU	Score	Calibration	Discrimination	
		Area Under ROC Curve (AUC)	Hosmer-Lemeshow Test	
			Chi-Square (8df)	Significance (p)
A	PIM	0.79	0,78	0,68
	PIM-2	0.84	3,81	0,149
	PRISM	0.91	3,43	0,329
	PRISM III	0.89	0,31	0,985
B	PIM	0.85	10,10	0,006
	PIM-2	0.90	1,94	0,379
	PRISM	0.89	0,19	0,667
	PRISM III	0.91	0,88	0,645
C	PIM	0.88	1,67	0,434
	PIM-2	0.93	12,06	0,006
	PRISM	0.84	2,79	0,425
	PRISM III	0.91	3,74	0,291

PRISM - Pediatric Risk of Mortality; PIM – Pediatric Index of Mortality; PICU – Pediatric Intensive Care Unit

Preliminary Analysis

- **Good general performance**

- Good discrimination power ($AUC > 0,8$);

- Well calibrated ($p > 0,05$);

- **Poor calibration of PIM2** ($p < 0,05$), although with a good discrimination power;

TABLE OF CONTENTS

- ▶ Introduction
 - ▶ Context
 - ▶ General Aims
- ▶ Methods
 - ▶ Study Design
 - ▶ Data Acquisition
 - ▶ Algorithms calculation
 - ▶ Statistical analysis
 - ▶ Discrimination
 - ▶ Calibration
- ▶ Results
 - ▶ Preliminary discussion
- ▶ **Project amplification**

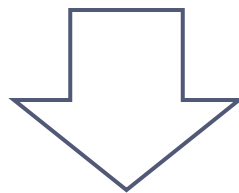
Optimization of PIM2 – a project amplification

Why?

- **Good** discriminatory power;
- **Poor** calibration;
- Bibliographic evidences identify PIM2 as showing **better performance** than PIM

How?

- **Binary Logistic Regression** with base on Portuguese data;
- **Re-estimation** of algorithm **coefficients** for Portuguese PICU's;



A optimized version for Portuguese reality of PIM2 score

Optimization of PIM2 – a project amplification

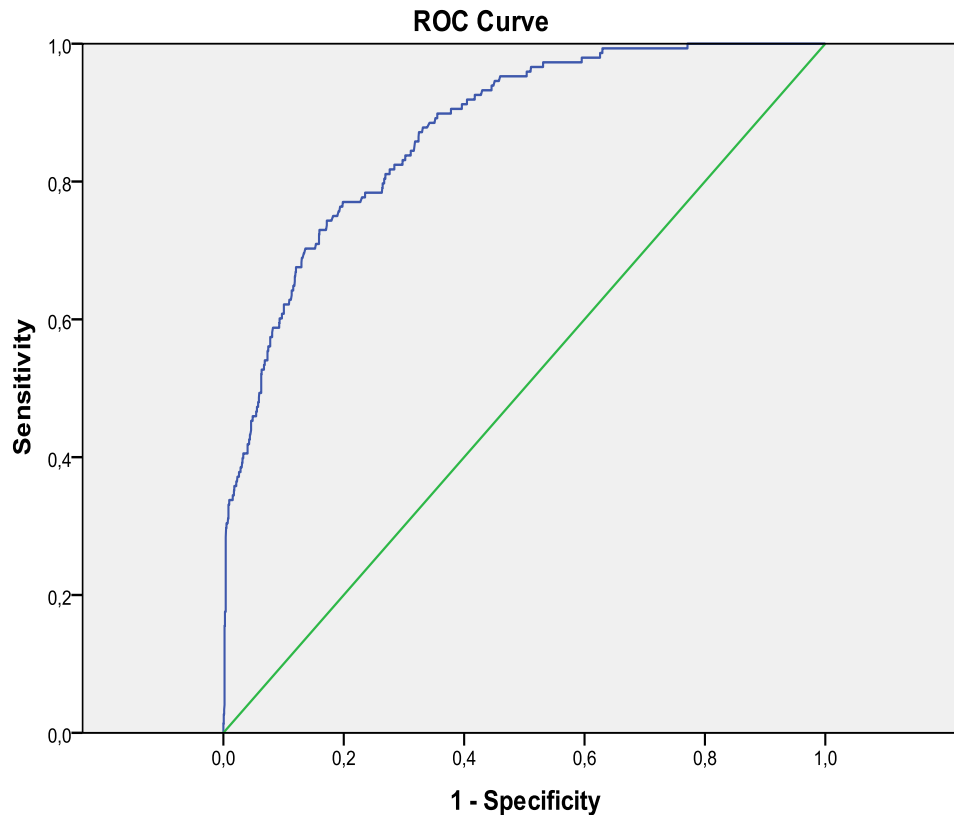
Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	PIM2_D	-18,230	40192,970	,000	1	1,000	,000
	PIM2_A	1,067	,216	24,338	1	,000	2,907
	PIM2_B	-1,298	,438	8,771	1	,003	,273
	PIM2_C	-2,055	,638	10,361	1	,001	,128
	PIM2_E	,654	,220	8,874	1	,003	1,923
	PIM2_F	-,008	,004	3,792	1	,051	,992
	PIM2_H	-,021	,019	1,131	1	,288	,980
	PIM2_I	3,200	,505	40,163	1	,000	24,532
	PIM2_J	-2,226	1,048	4,514	1	,034	,108
	PIM2_GRec	-,046	,329	,020	1	,888	,955
	Constant	-1,926	,455	17,918	1	,000	,146

a. Variable(s) entered on step 1: PIM2_D, PIM2_A, PIM2_B, PIM2_C, PIM2_E, PIM2_F, PIM2_H, PIM2_I, PIM2_J, PIM2_GRec.

Re-estimated coefficients for optimized version construction

Optimized portuguese PIM2 version discrimination analysis



Diagonal segments are produced by ties.

Area Under the Curve

Test Result Variable(s): Predicted Probability PIM2 T4 (%)

Area	Asymptotic 95% Confidence Interval	
	Lower Bound	Upper Bound
,872	,845	,900

PIM – Pediatric Index of Mortality;

PROJECT AMPLIFICATION

Optimized portuguese PIM2 version calibration analysis

Contingency Table for Hosmer and Lemeshow Test

		Estado clínico na alta da UCIP = Vivo		Estado clínico na alta da UCIP = Falecido		Total
		Observed	Expected	Observed	Expected	
Step 1	1	561	559,793	1	2,207	562
	2	257	258,262	6	4,738	263
	3	589	592,420	54	50,580	643
	4	135	131,524	87	90,476	222

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	1,481	2	,477

PIM – Pediatric Index of Mortality;

ACKNOWLEDGEMENTS

- ▶ The authors gratefully acknowledge the guidance provided by professors Armando Teixeira Pinto and Rosa Oliveira.

REFERENCES

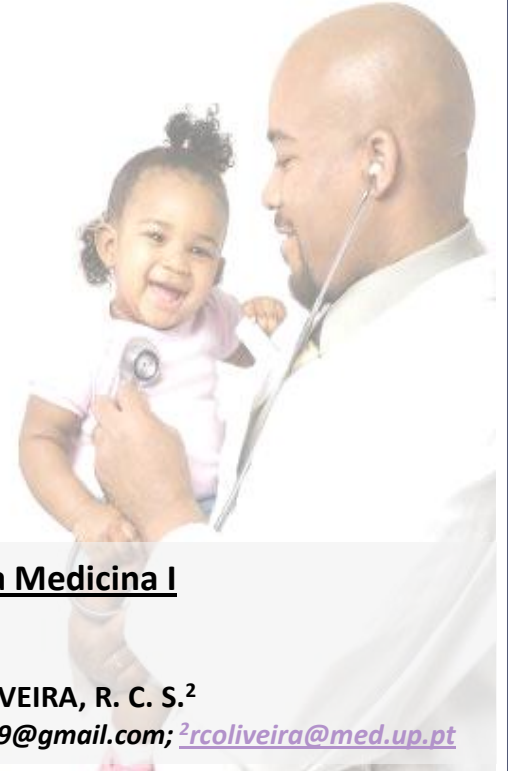
- ❖ ¹Gemke RJ, Bonsel GJ, Bught AJ. Outcome assessment and quality assurance in pediatric intensive care. In: Tibboel D, van der Voort E, editors. Intensive care in childhood – a challenge to future. 2nd ed. Berlin: Springer; 1996. p. 117-32.
- ❖ ²Gunning K, Rowan K. ABC of intensive care outcome data and scoring systems. *BMJ*. 1999;319:241-4.
- ❖ ³ Reliability. (2009). In *Merriam-Webster Online Dictionary*. Retrieved December 1, 2009, from, <http://www.merriam-webster.com/dictionary/reliability>
- ❖ ⁴ Messick, S. (1989). *Validity*. In R. L. Linn (Ed.), *Educational measurement* (pp. 13-103). New York: American Council on Education/MacMillan.
- ❖ ⁵ Brady, R. A. (2006). *Assessment and Optimization of Mortality Prediction Tools for Admissions to Pediatric Intensive Care in the United Kingdom*. *Pediatrics*;
- ❖ ⁶ Francisco, G. A. (2003). *Performance of SAPS3, compared with APACHE II and SOFA, to predict hospital mortality in a general ICU in Portugal*; *European Journal of Anaesthesiology*;
- ❖ ⁷ Harrell FE, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA*. 1982;247:2543–2546
- ❖ ⁸ Murphy-Filkins R, Teres D, Lemeshow S, Hosmer DW. Effect of changing patient mix on the performance of an intensive care unit severity-of-illness model: how to distinguish a general from a specialty intensive care unit. *Crit Care Med*. 1996;24:1968–1973
- ❖ ⁹ Deeks, J. J.; Altman, D. G. (2004); *Diagnostic tests 4: likelihood ratios* ; *Intensive Care Med*;

Performance of Mortality prognostic scores in Portuguese PICU's

“Are they doing a good job?”

THANK YOU!

ANY QUESTIONS?



Introdução à Medicina I

CLASS 4¹

Adviser: OLIVEIRA, R. C. S.²

¹ turma4fmup09@gmail.com; ² rcoliveira@med.up.pt