

Machine learning for health

-

Examples & discussion



Xavier Tannier



Attribution-NonCommercial-ShareAlike

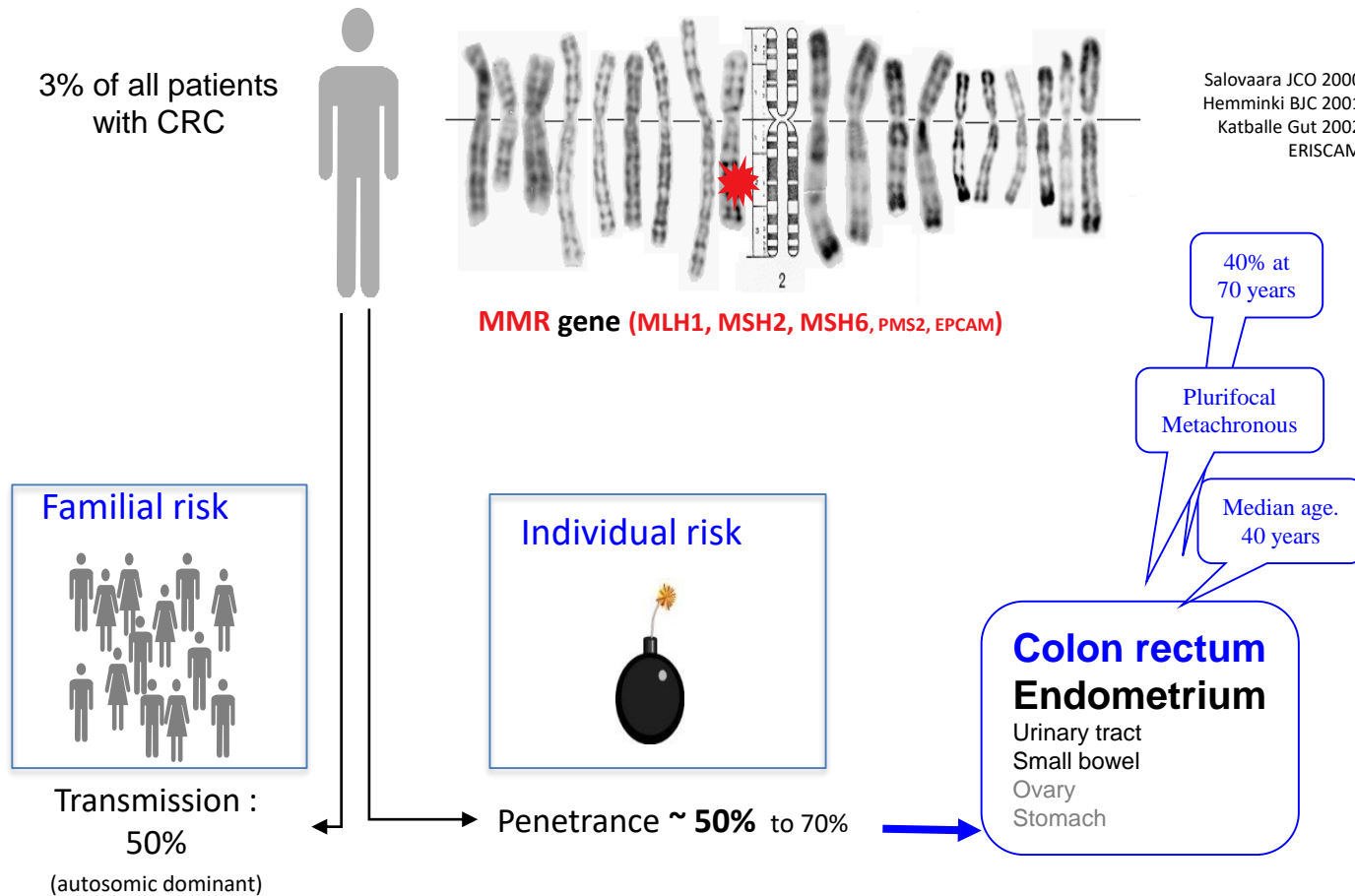
CC BY-NC-SA



2023

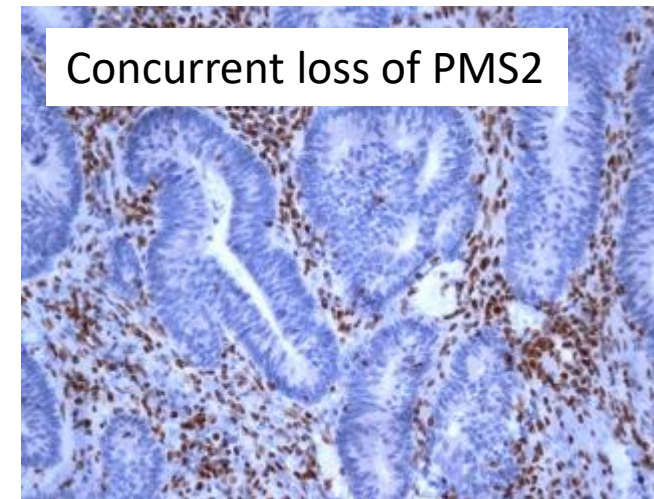
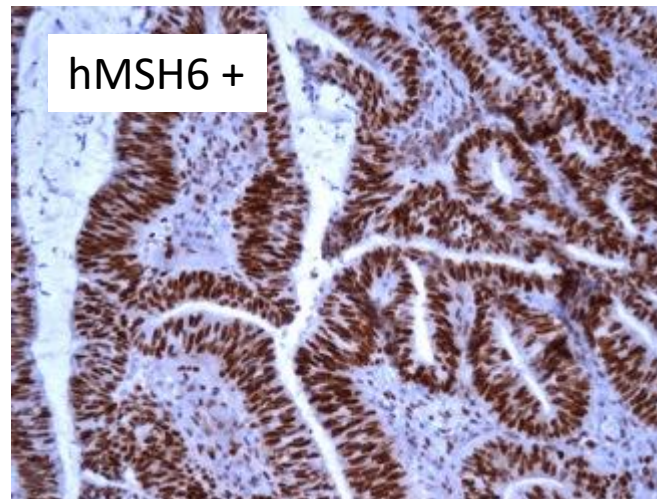
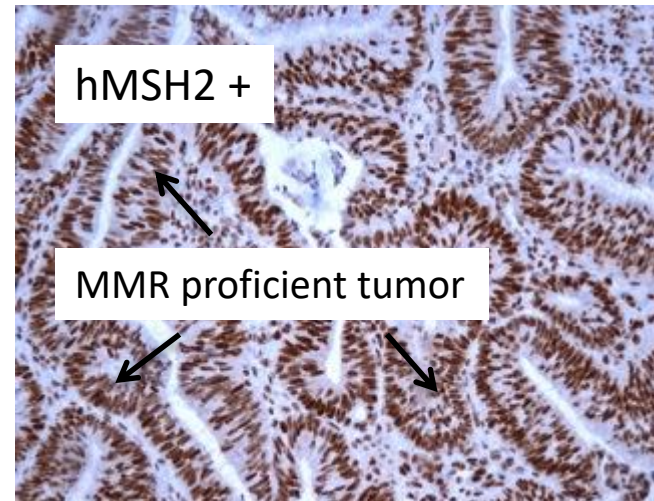
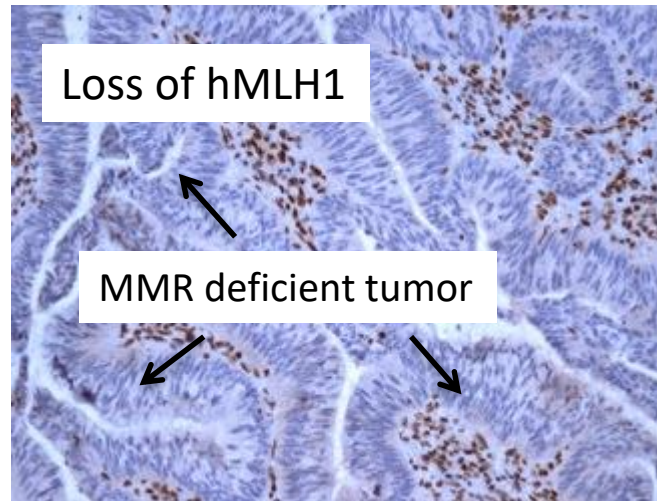
Lynch syndrome diagnosis

Lynch syndrome, the most common genetic predisposition to cancer, including colorectal cancer



Immunohistochemistry

Unstable tumor (MSI): loss of MMR protein expression



Cas personnel F. Bibeau

Recent approach

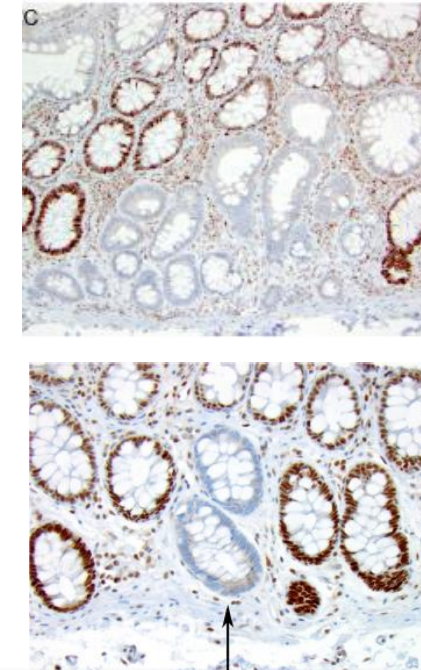
DNA mismatch repair protein deficient non-neoplastic colonic crypts: a novel indicator of Lynch syndrome

Rish K. Pai¹ · Beth Dudley² · Eve Karloski² · Randall E. Brand² · Neil O'Callaghan^{3,4} · Christophe Rosty^{3,4,5,6} · Daniel D. Buchanan^{3,4,7} · Mark A. Jenkins⁸ · Stephen N. Thibodeau⁹ · Amy J. French⁹ · Noralane M. Lindor¹⁰ · Reetesh K. Pai¹¹

Aim: To determine if MMR-deficient non-neoplastic intestinal crypts can help distinguish between patients with and without Lynch Syndrome

Table 1 Clinical and pathologic features of patients with and without Lynch syndrome

Clinical and pathologic features	Confirmed Lynch syndrome <i>N</i> (%)	Non-Lynch syndrome				<i>p</i> value***
		Overall <i>N</i> (%)	Lynch-like syndrome** <i>N</i> (%)	Sporadic MLH1 protein deficiency <i>N</i> (%)	MMR proficient <i>N</i> (%)	
No. of cases	52	70	10	30	30	NA
MMR protein expression of affected MMR gene in normal intestinal crypts*						
Intact expression	34 (65)	69 (99)	9 (90)	30 (100)	30 (100)	<0.001
Loss of expression	18 (35)	1 (1)	1 (10)	0 (0)	0 (0)	
MLH1	4	0	0			
MSH2	7	1	1			
MSH6	4	0	0			
PMS2	3	0	0			

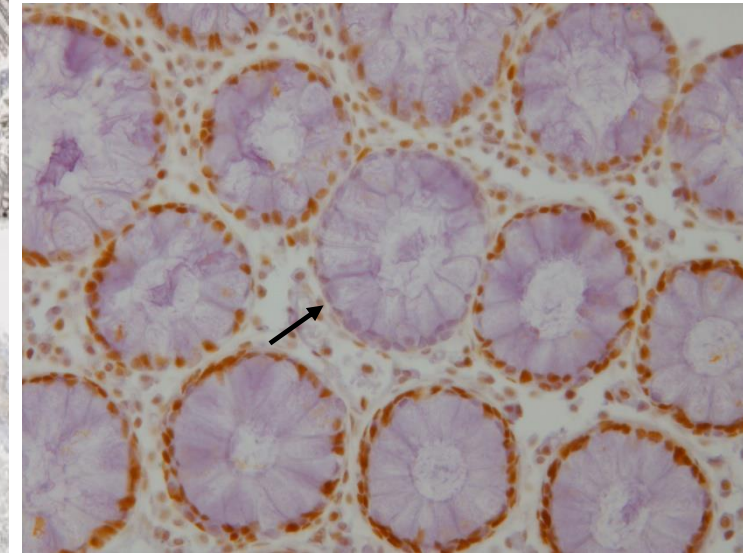
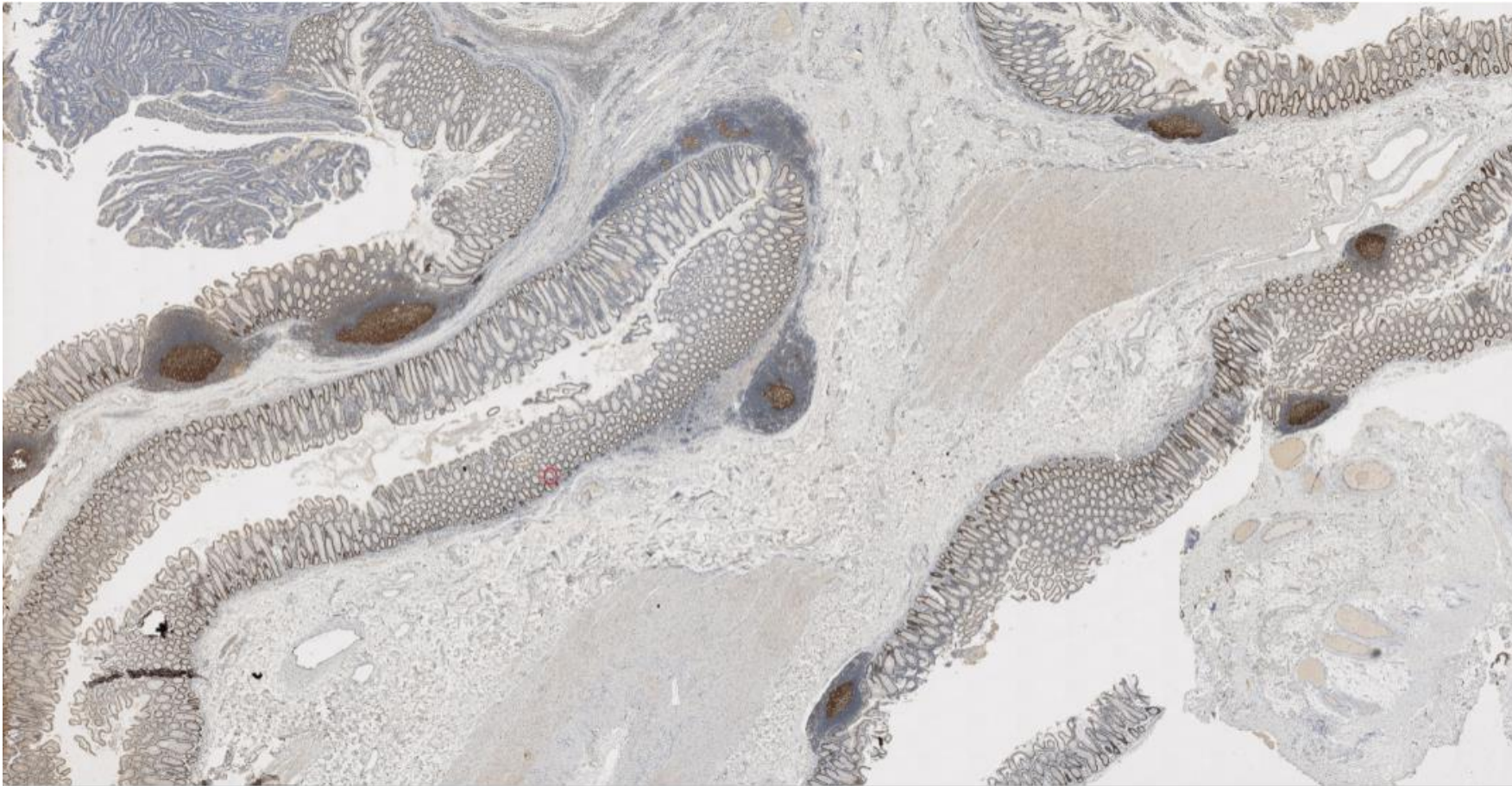


MMR-deficient crypts are precursor lesions that are seen in the non-tumor colonic mucosa only in patients with Lynch syndrome.

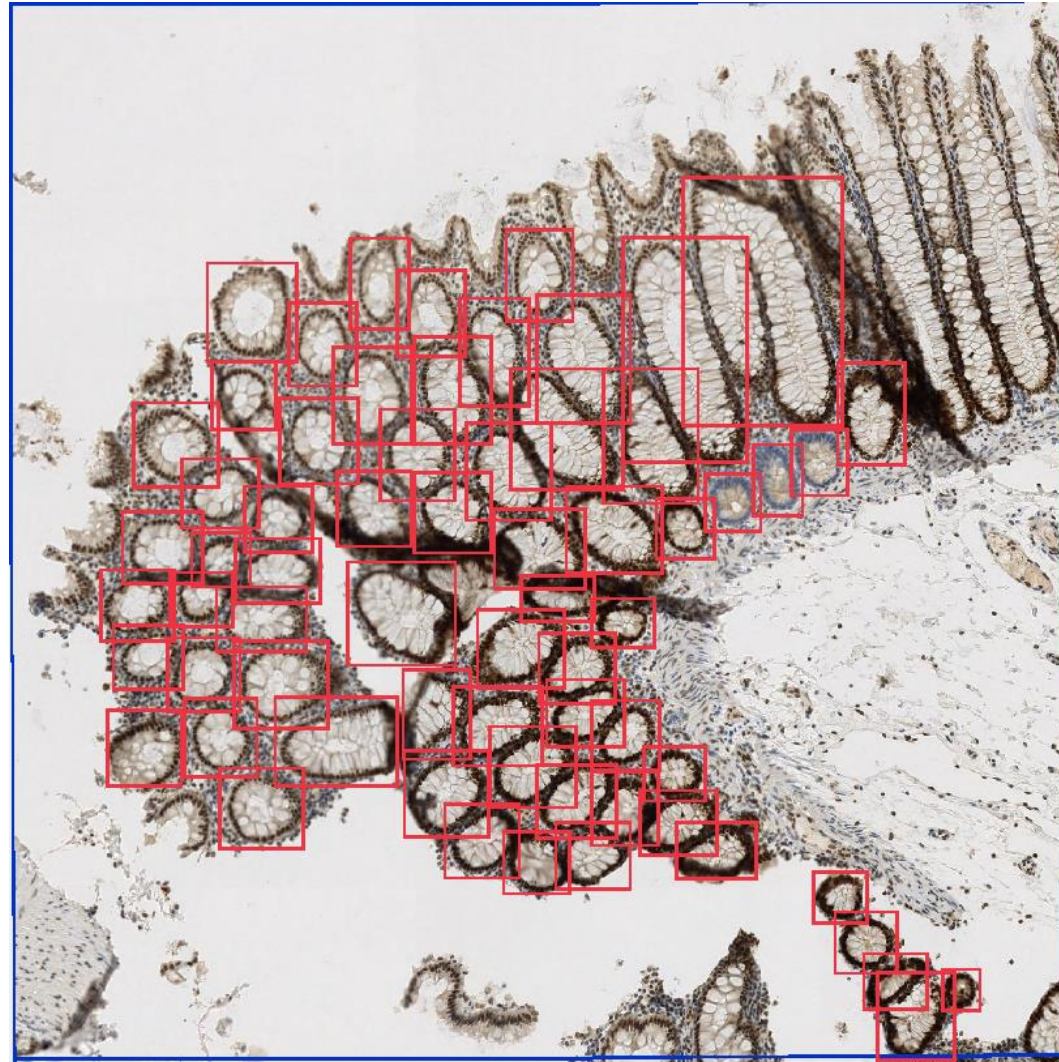
The evaluation of these lesions can therefore be an aid to the diagnosis of Lynch syndrome when constitutional testing is not contributory.

Searching for MMR-deficient crypts, a long and tedious diagnostic approach

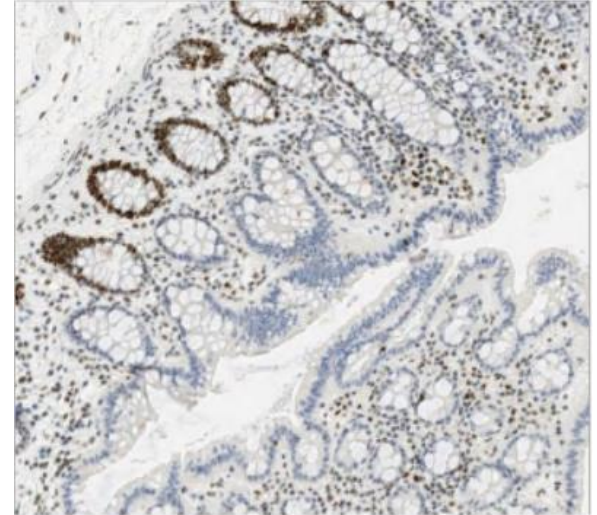
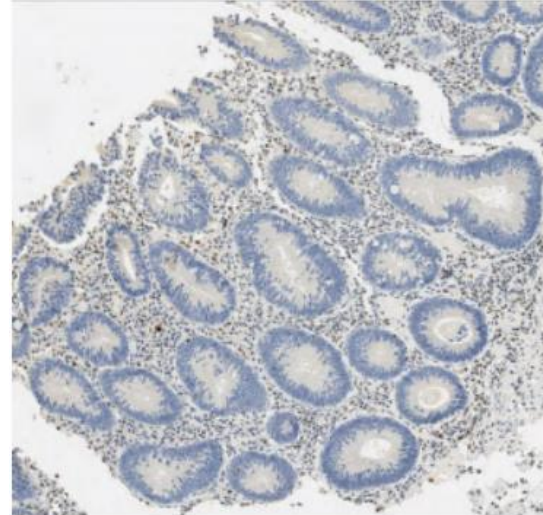
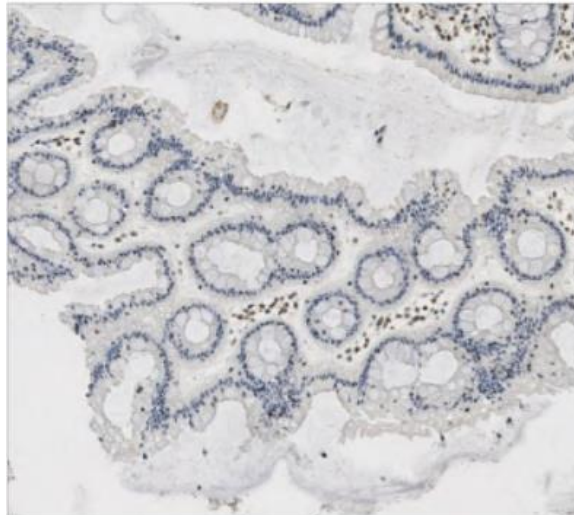
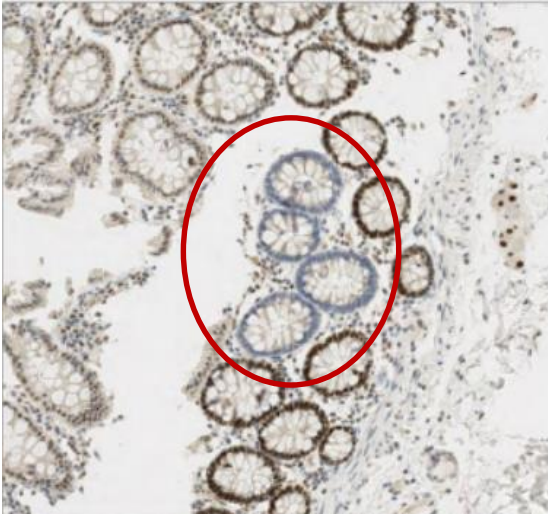
95% probability of identifying at least one MMR-deficient crypt if analysis of 26.3 mm² of normal colonic mucosa;
need to evaluate 10 immunohistochemistry slides (Pai et al., Modern Pathol 2018)



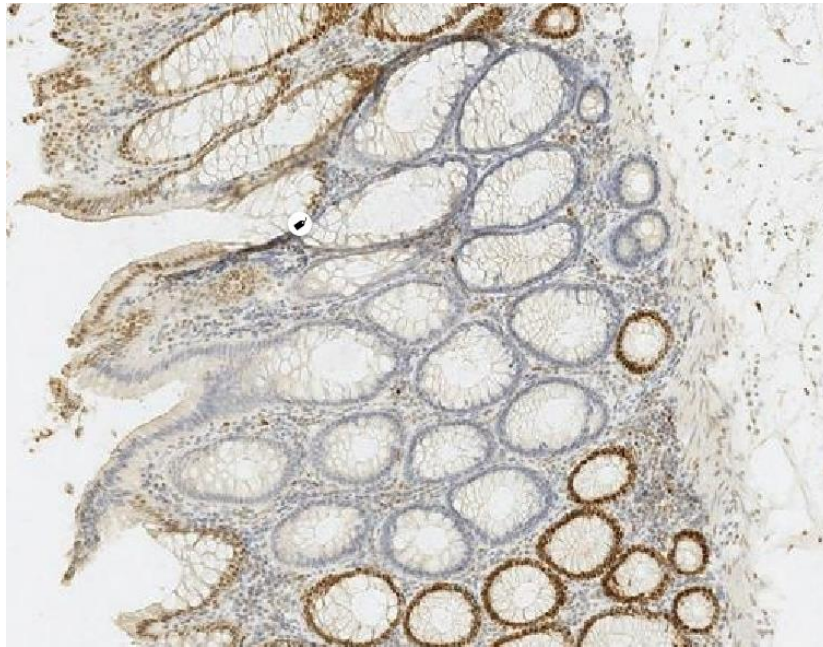
Crypt detection



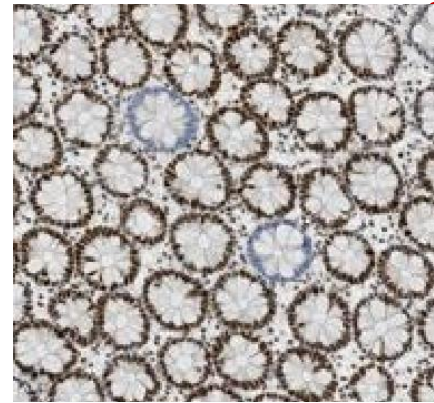
Classes are ambiguous



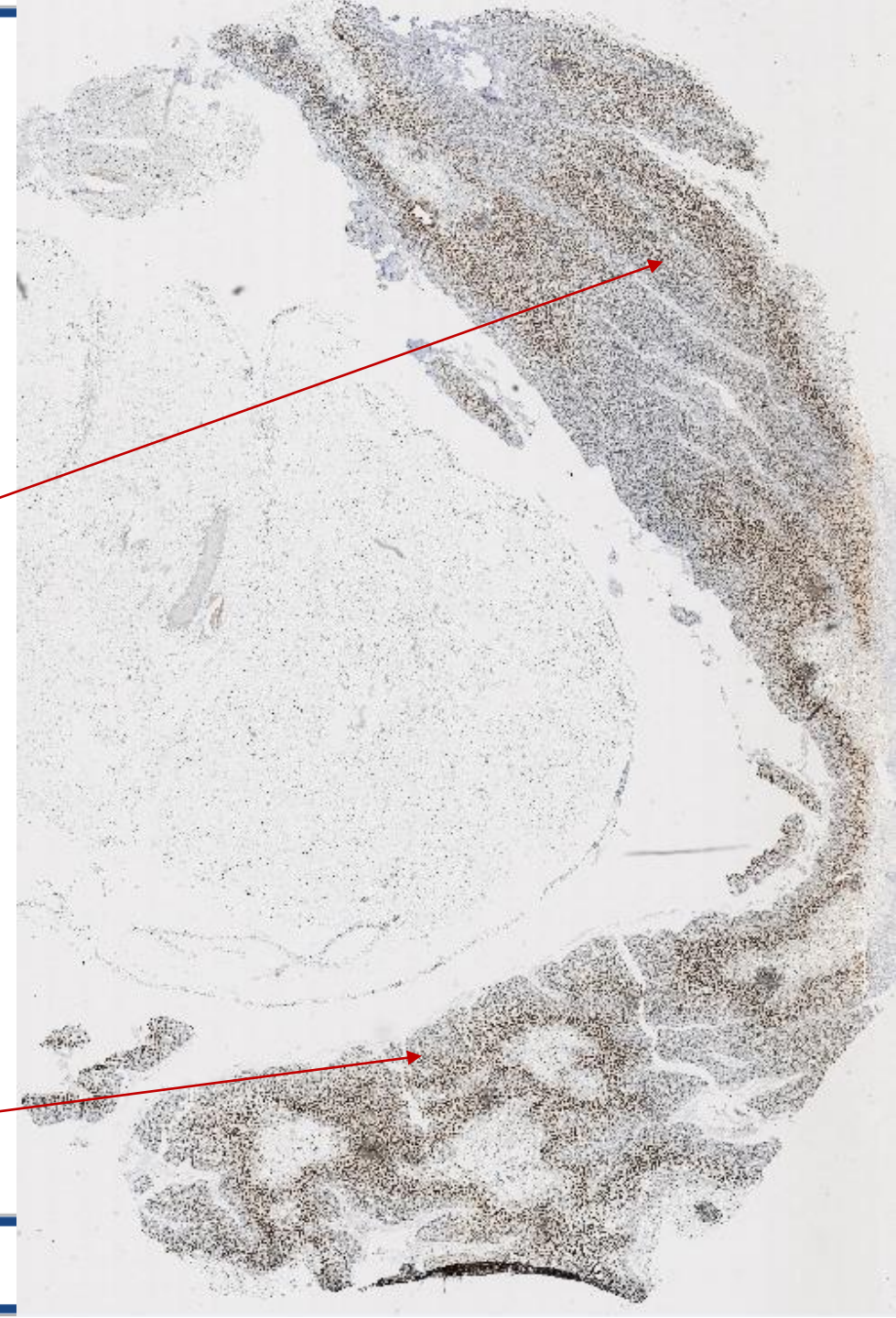
Classes are ambiguous



Skewed distribution



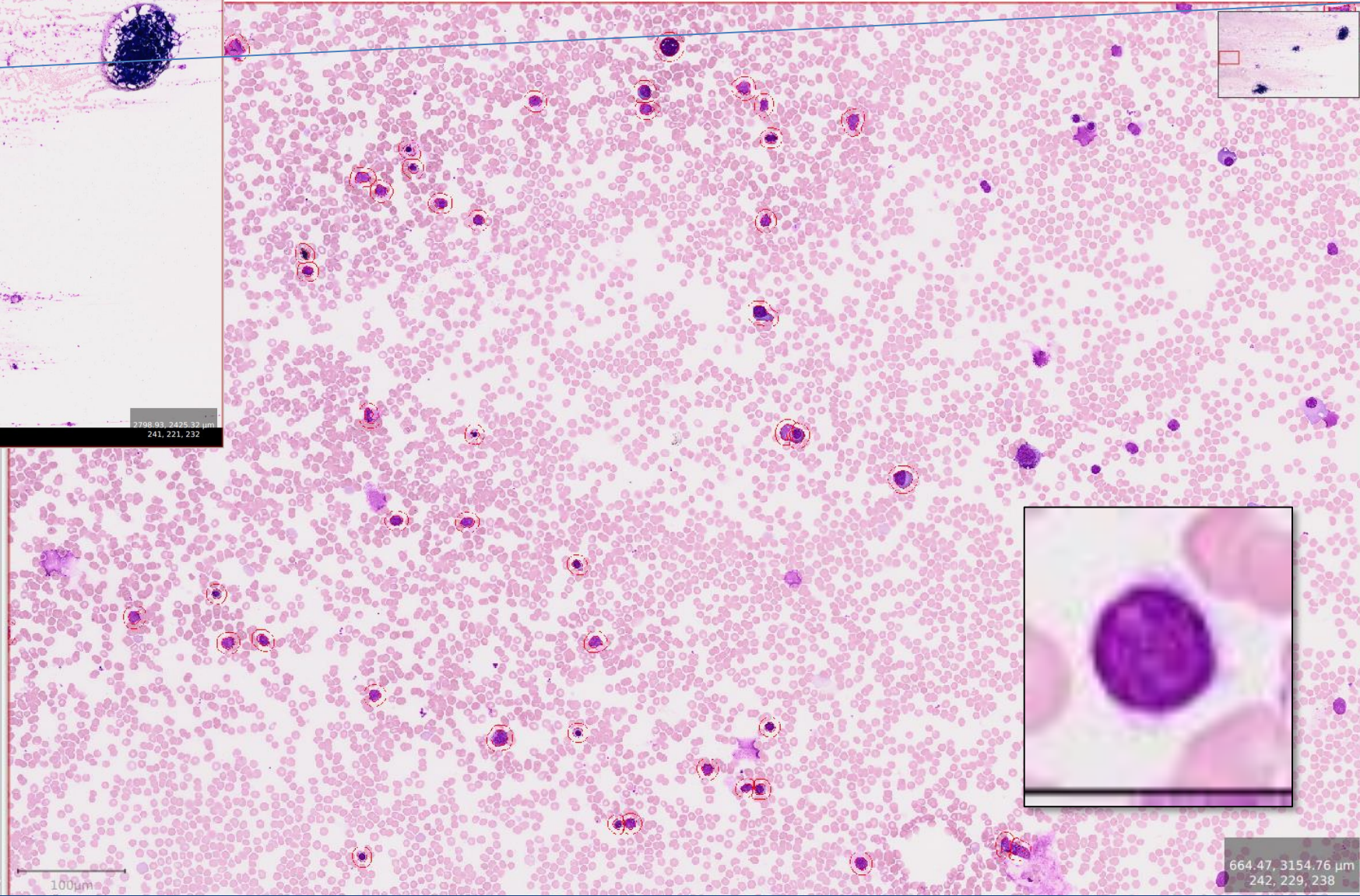
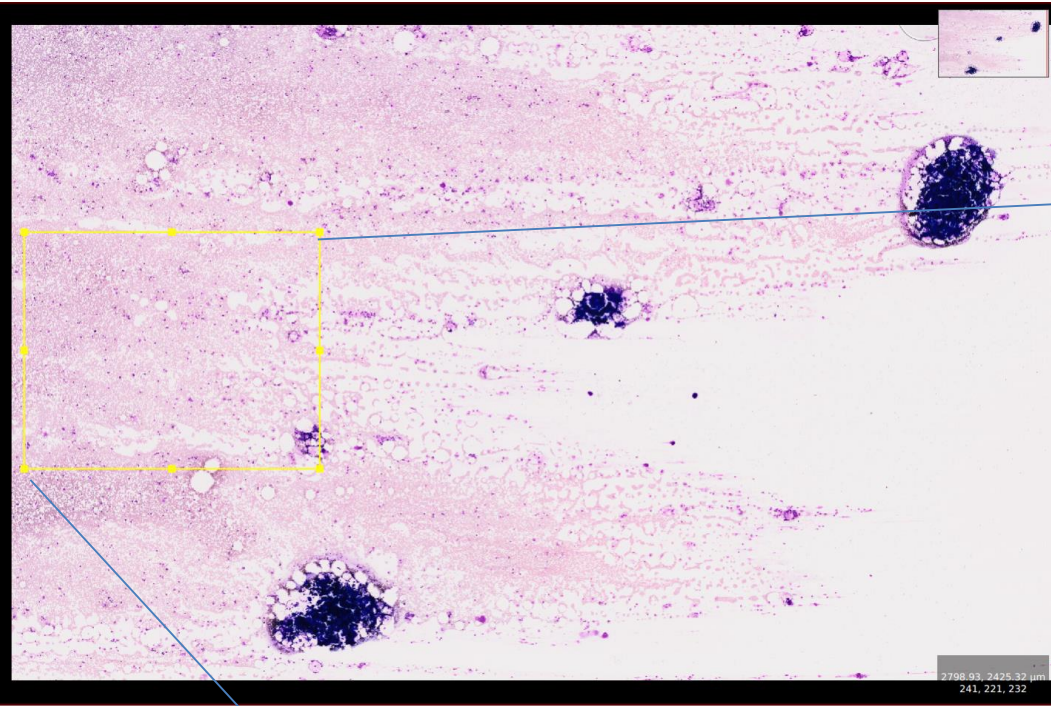
2023



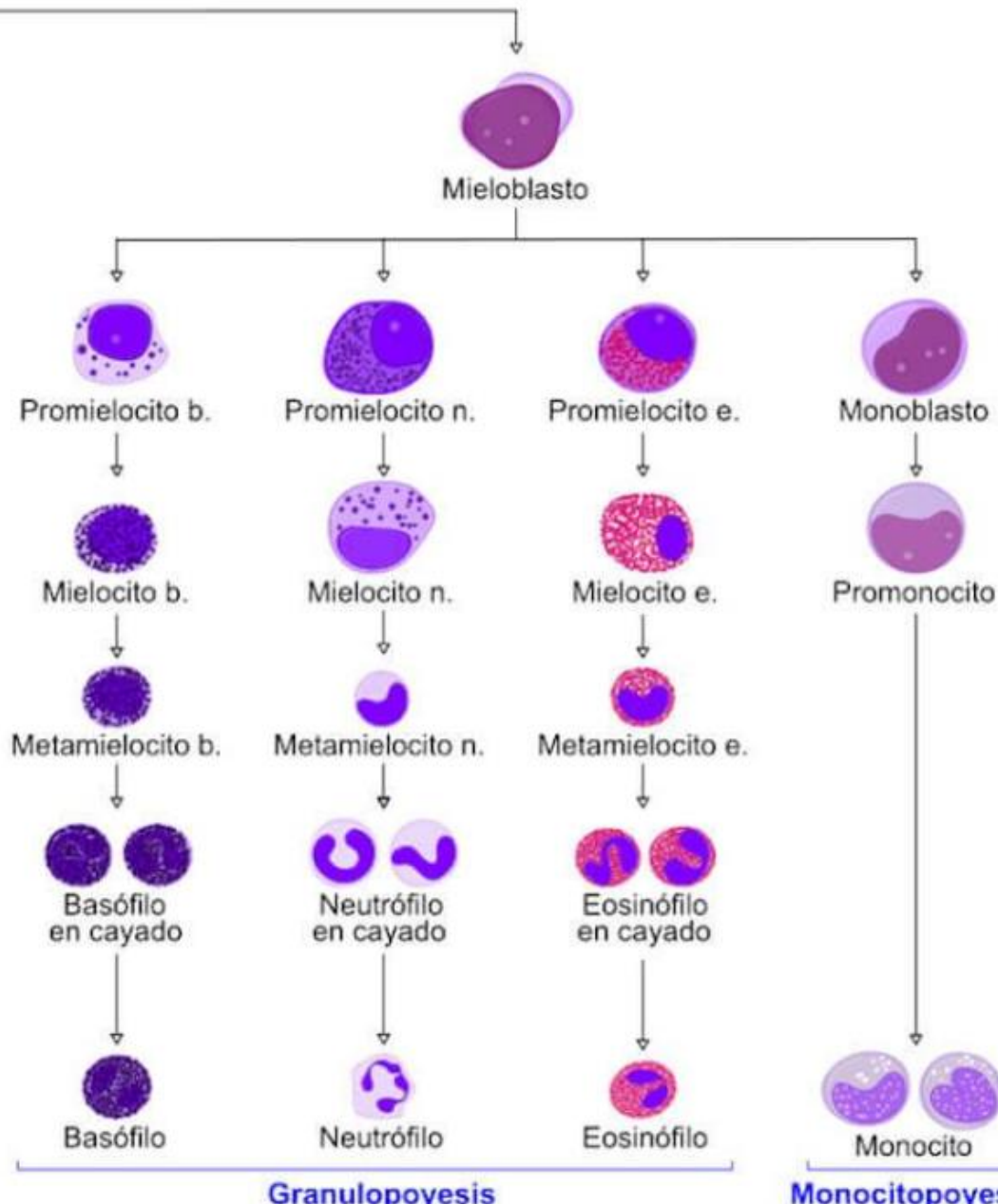
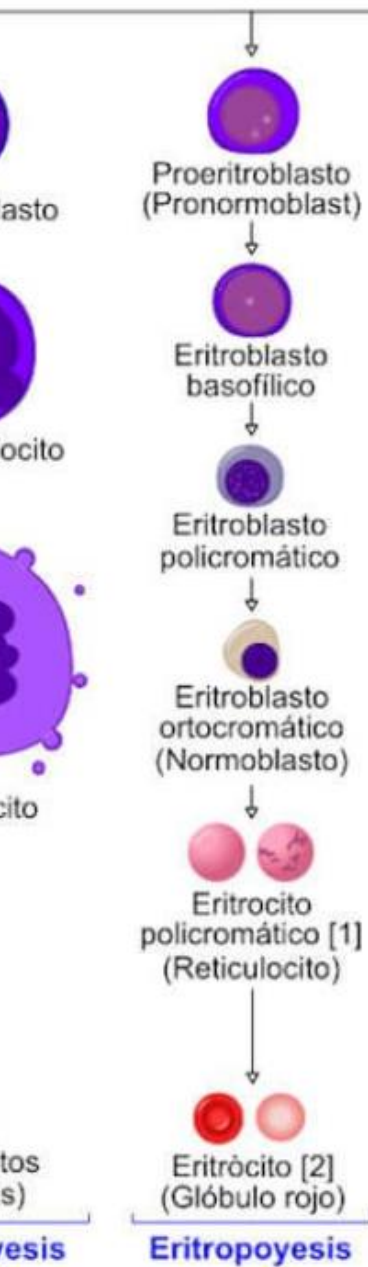
White blood cell classification

a teaser of the practical sessions with Manon Chossegros

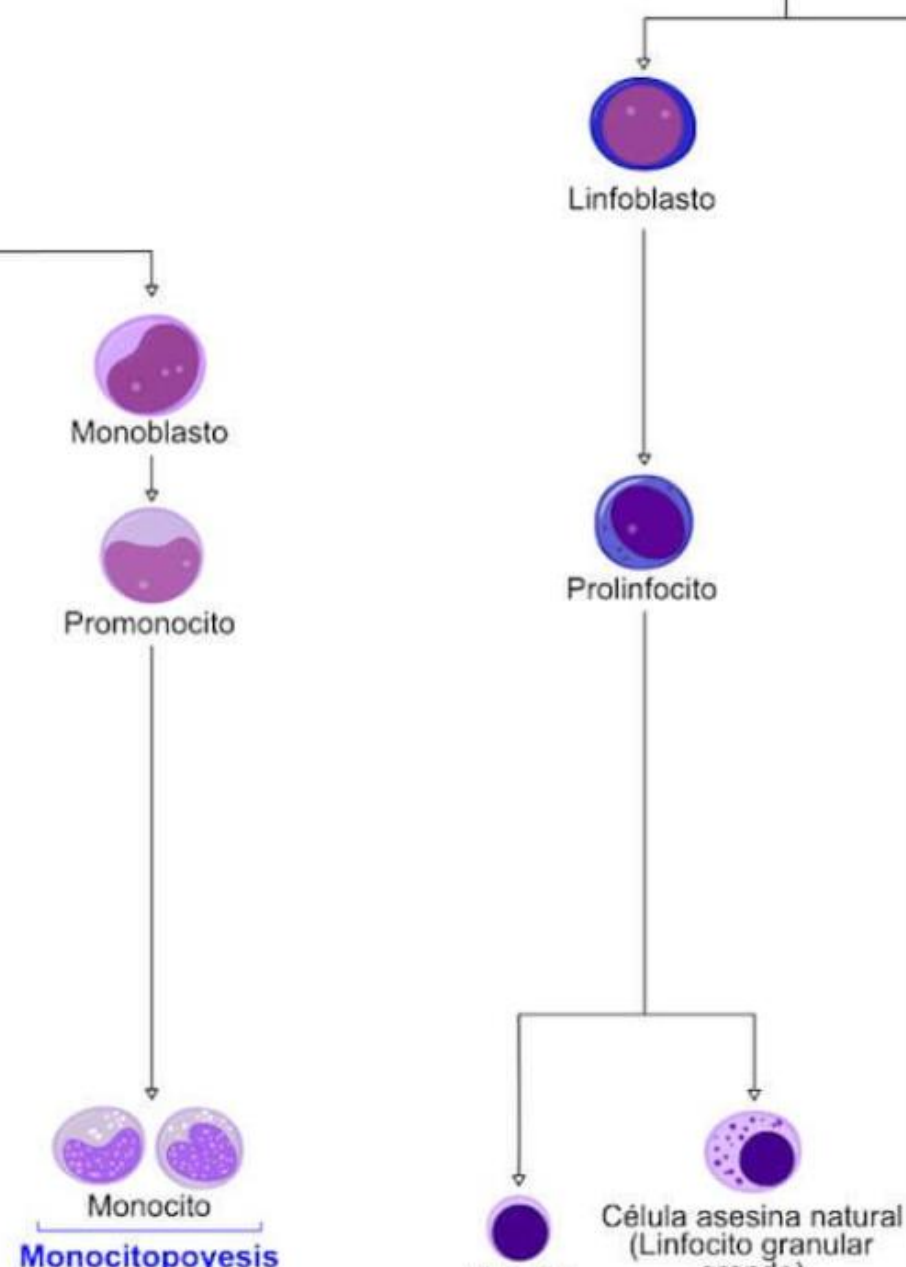
Blood smear



précurseur commun mieloïde



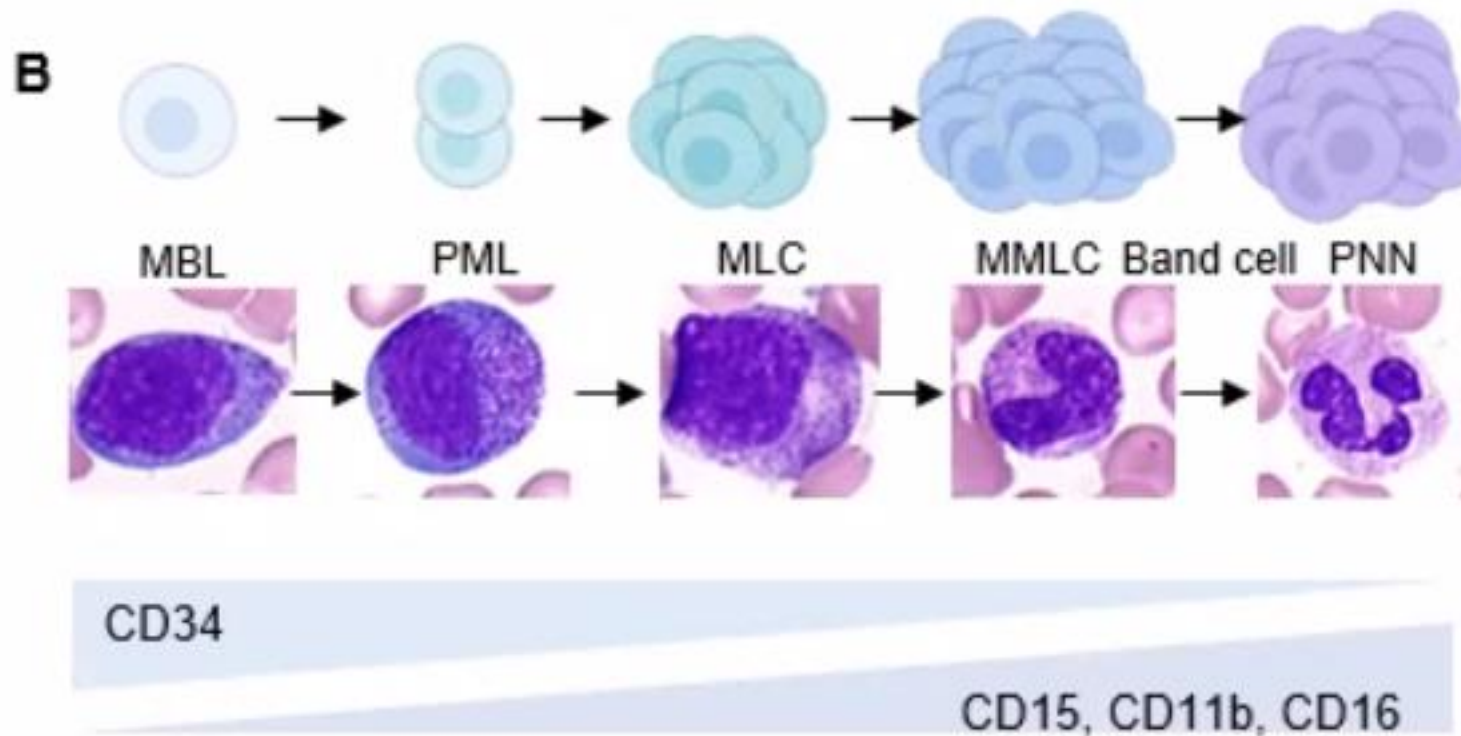
précurseur commun lymphoïde



Células precursoras

Armas

White blood cell evolution

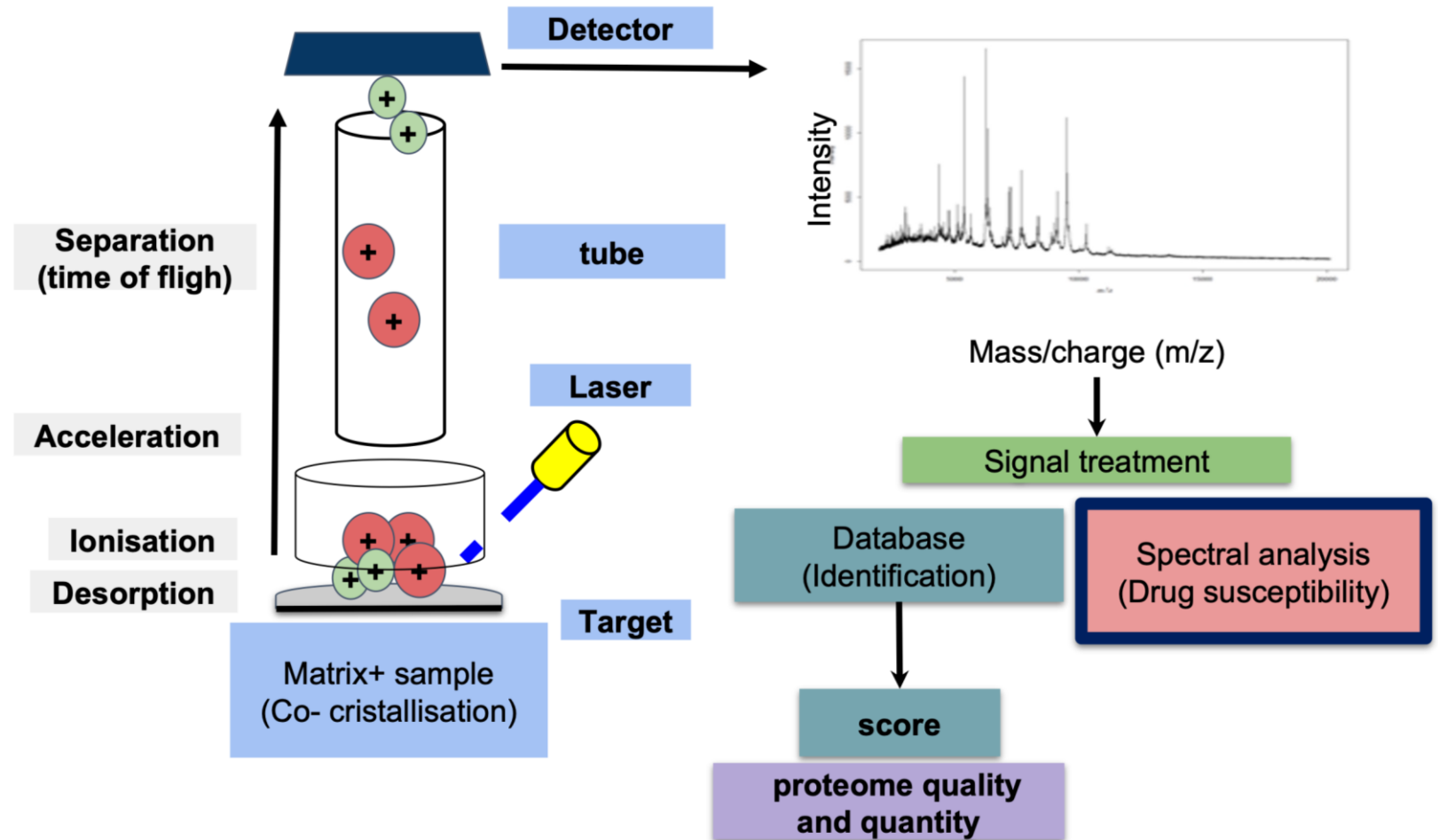


Mass spectrometry

Maldi TOF mass spectrometry

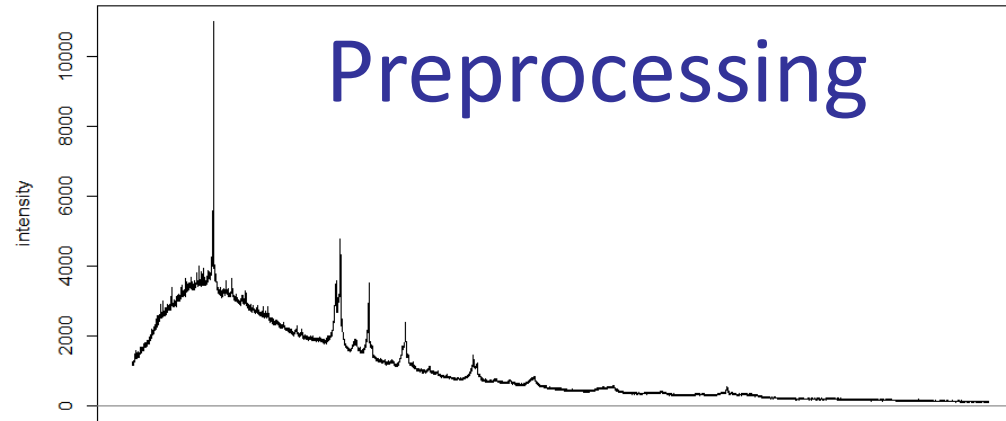


Spectrometer
Microflex LT Bruker®

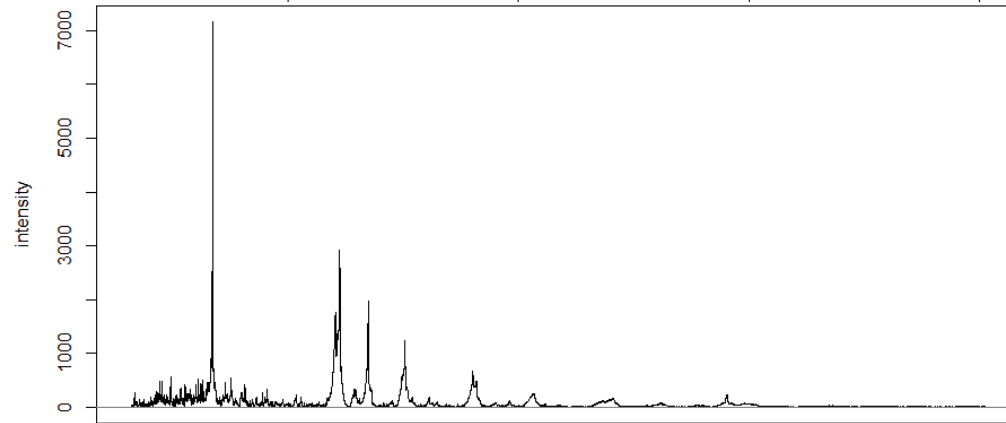


Preprocessing

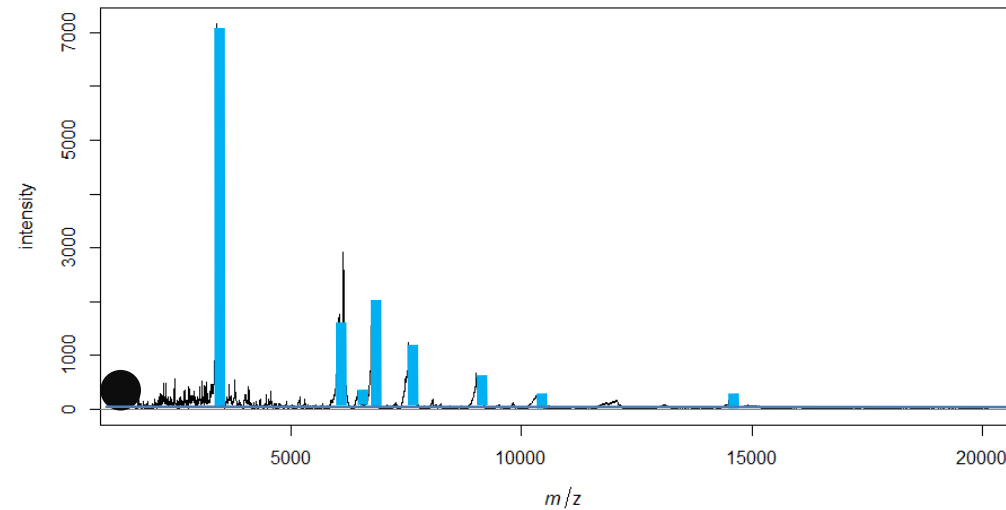
Raw spectrum



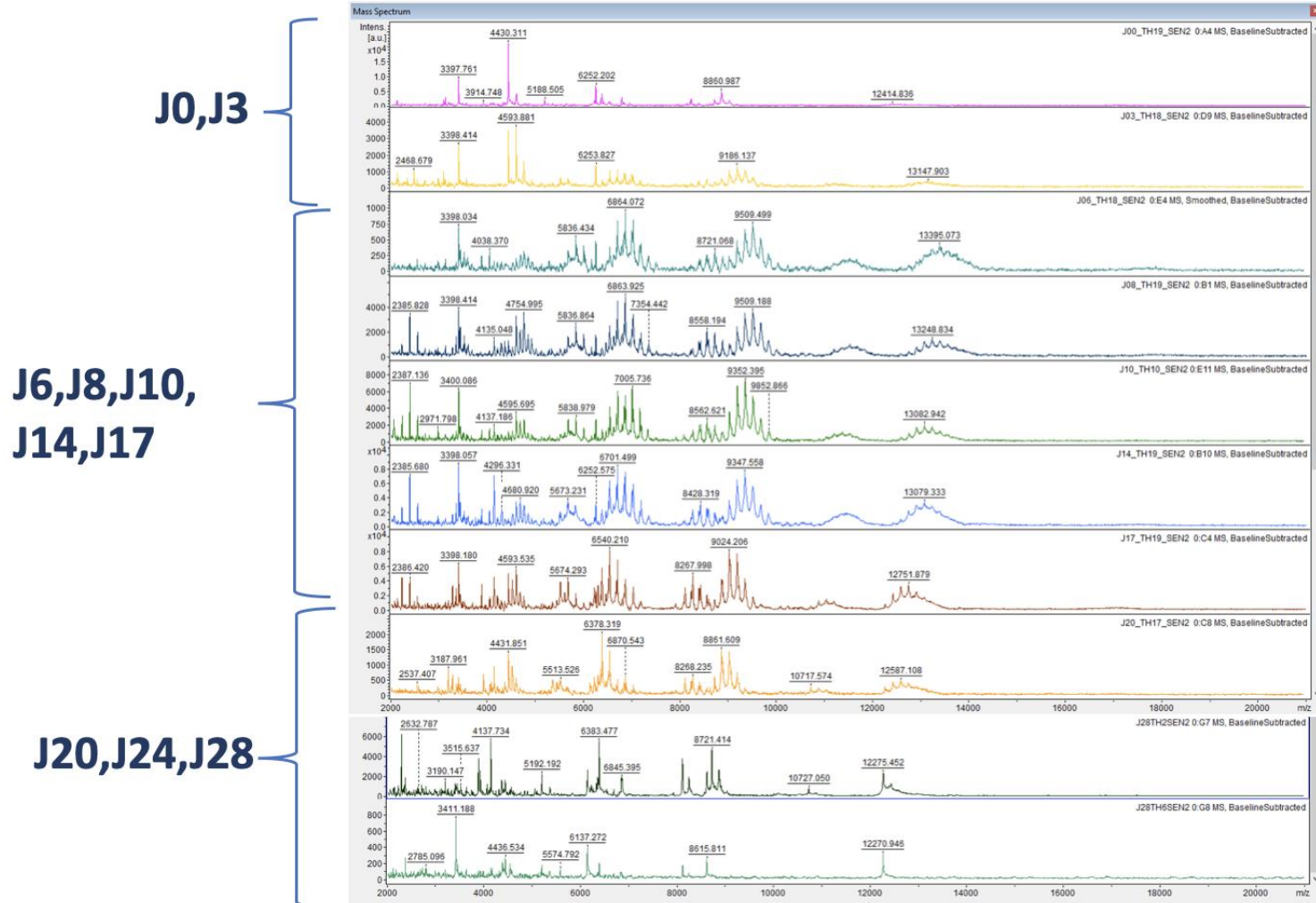
Baseline correction



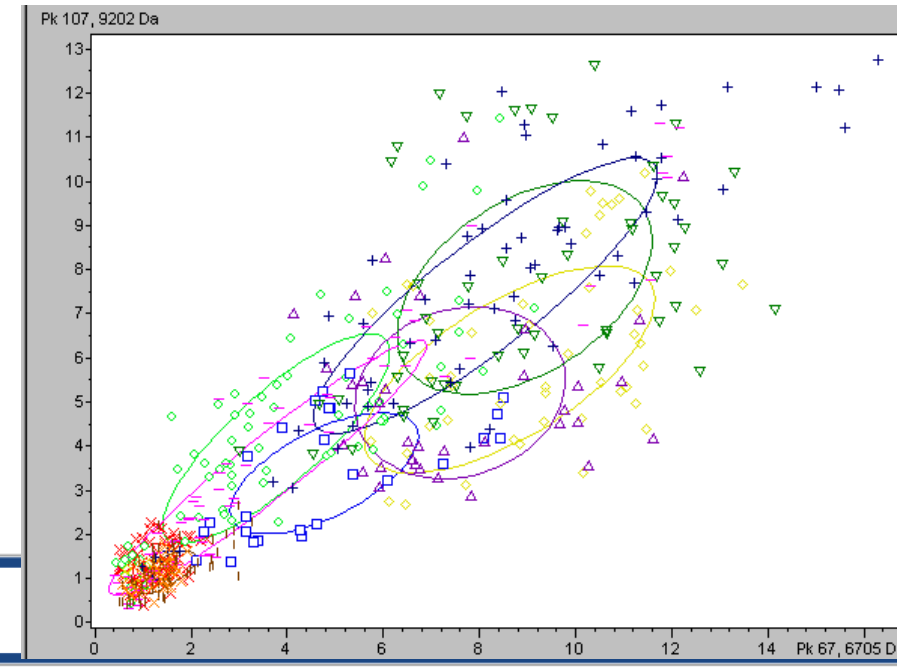
Peak identification



Visualization of spectral profiles (the example of mosquito age identification)



Protein profiling
(ClinProTool)

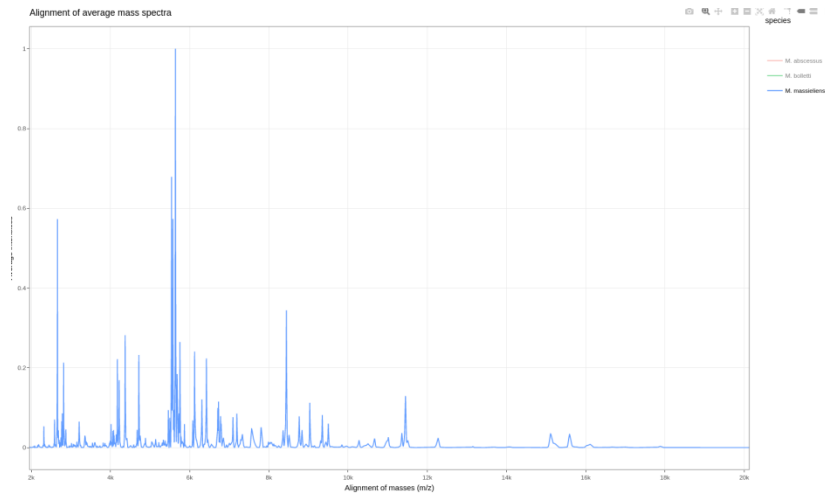


96 *Candida parapsilosis* isolates typed by the microsatellite approach. Our task is to identify the clone isolates (in blue).

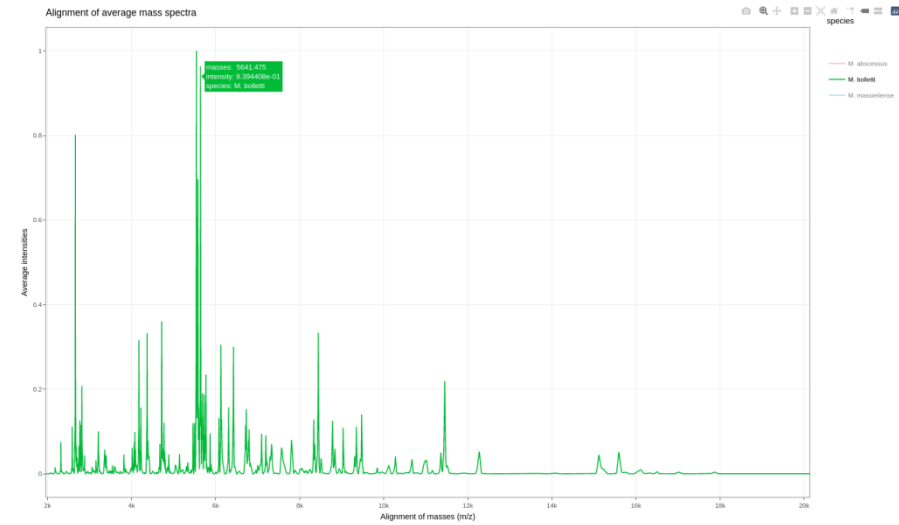


Identification of subspecies

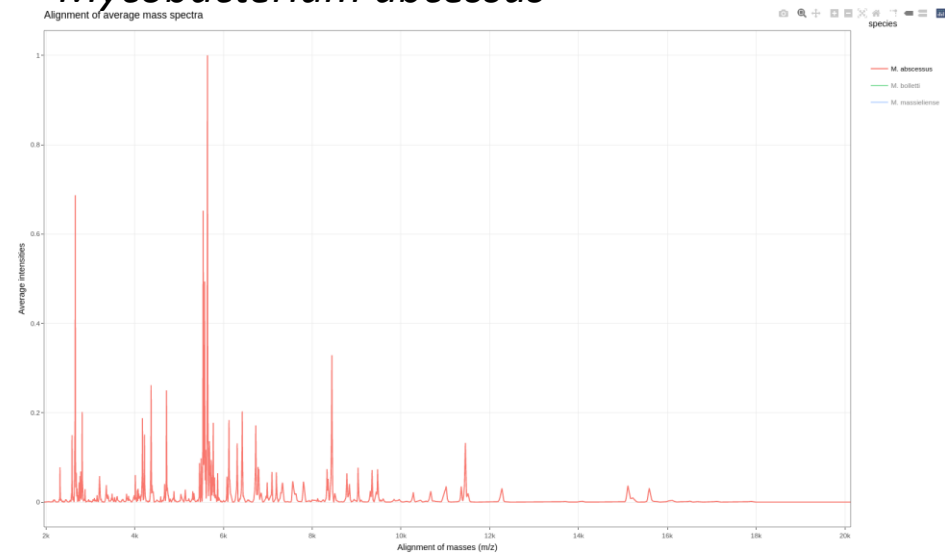
Mycobacterium massiliense



Mycobacterium bolletii

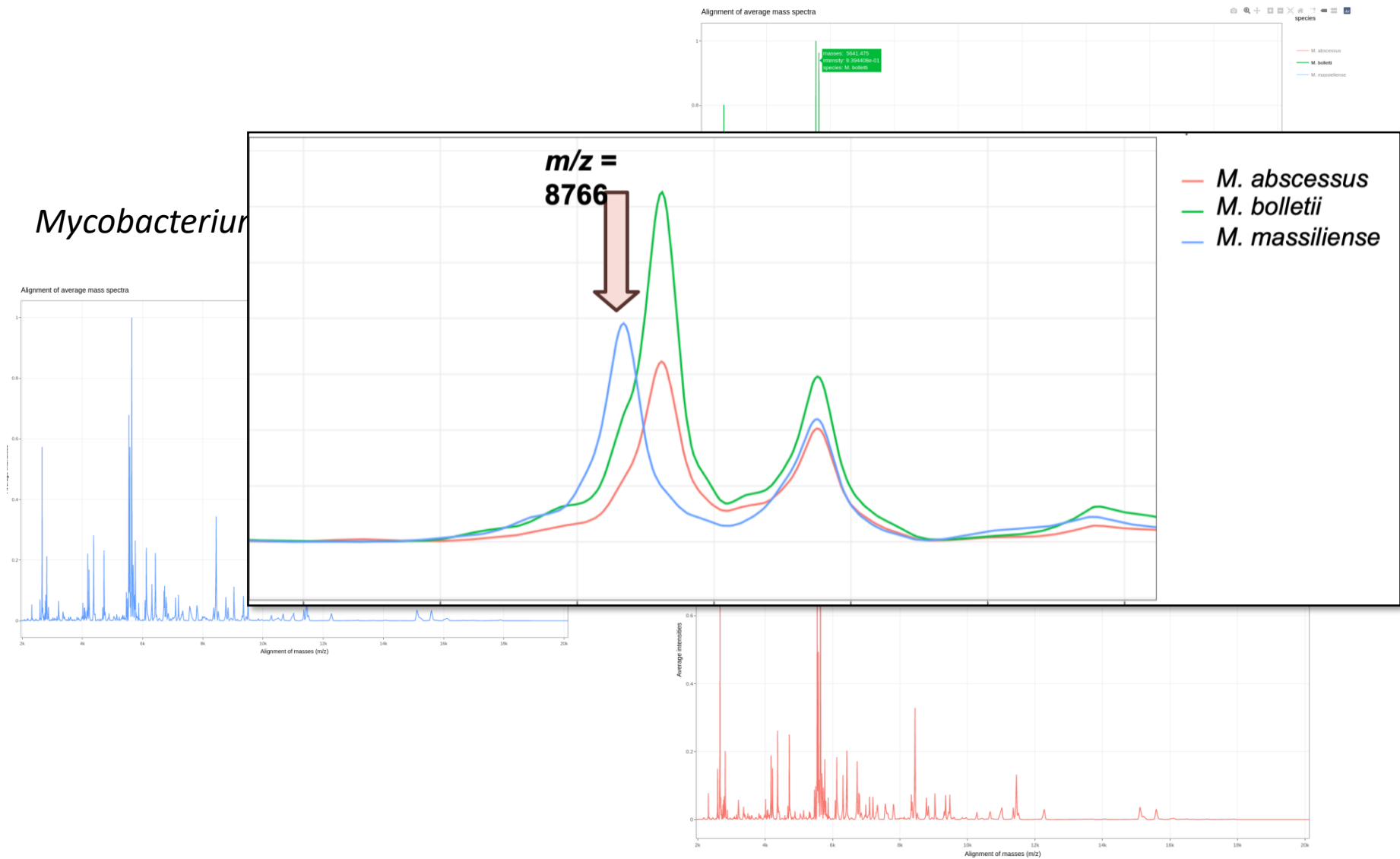


Mycobacterium abscessus



Identification of subspecies


Mycobacterium bolletii



NLP tasks for health documents

a teaser of the practical sessions with Christel Gérardin

Different tasks

- 
1. Concept extraction
 2. Concept normalization
 3. Relation extraction
 4. Document classification
 5. Patient classification

- The last steps generally involve the first ones.
- Each step should be evaluated

Concept extraction and characterization

CONCLUSION_COVID [diagnosis:probable]

Pulmonary lesion very evocative of Covid

PROBLEM [fact:neg]

Patient with no evident sign of cranial traumatism

PROCEDURE [fact:patient history]

Patient with history of bariatric surgery

9 blood cultures the 26/6/15 to
Staphylococcus aureus methicilline sensitive.

BACTERIA RESISTANCE

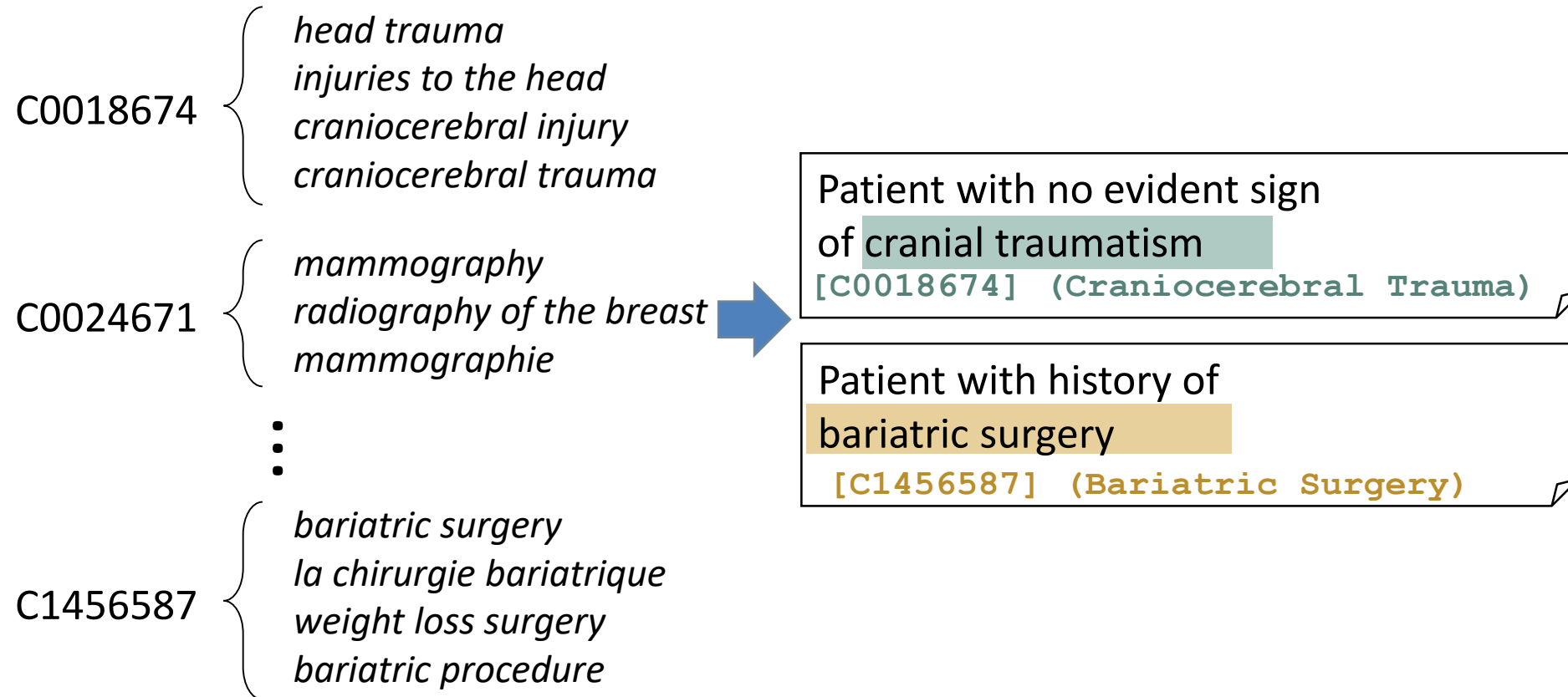
BACTERIA

Staphylococcus aureus
Escherichia coli
Streptococcus pneumoniae
Streptococcus pyogenes
Klebsiella pneumoniae

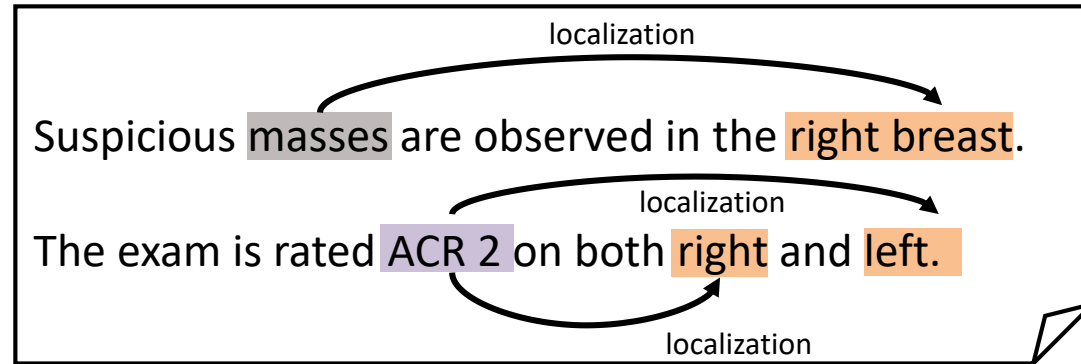
BACTERIA
RESISTANCE

<BACTERIA> + . * sensitive
<BACTERIA> + . * resistant
<BACTERIA> + sensitive to methicillin
...

Concept normalization



Relation extraction



Phenotyping: text/patient classification

[illegible]

Reason for admission =
low kinetic fracture ?

Family history?

ASSISTANCE HÔPITALIAIRE

ASSISTANCE HÔPITALIAIRE

ASSISTANCE HÔPITALIAIRE
PUBLIQUE DE PARIS
HÔTEL - DE DIEU

HÔTEL DIEU
101 rue de la Harpe, Paris

Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

Télégramme
HÔTEL DIEU PARIS
Télégramme
HÔTEL DIEU PARIS

N° de ...

Prénoms ...

Date Naissance ...

N° ...

EV

SCIENTIFIQUE AU GALLIUM 67

Demandeur : Médecin de l'Unité 7055

Docteur : EDWARD-CHÉRIE DEB

N° de Demande : 1070473

Date de Demande : 01/01/1967

Date d'Examen : 01/01/1967

Paris, le 01/01/1967

CLINIQUE :
Médicalisation à macroscopie, dans un contexte d'aggravation nette des douleurs et de
Dépression fonctionnelle.

Possibilité chirurgicale : Éléments importants dans les caractéristiques de la maladie associée à des
signes de gravité.

TÉCHNIQUE : Examen réalisé 48 heures après l'ingestion de 18,5 MmBz de Chlorure de Gallium 67
après 10 jours de traitement au corticoïde. Synthèse tumorale.

RÉSULTAT :

Les clichés réalisés sur la scintigraphie suscitent une bonne délimitation de l'articulation du
quadrilatère qui apparaît diffuse et plus intense homogène métraculaire au versant des articulations
de la moelle sus-jacente.

Les clichés réalisés sur la scintigraphie suscitent une bonne délimitation de l'articulation du
quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin.

Sur la partie inférieure des clichés et sur les radiels, on note des images scintigraphiques homogènes plus
intenses, bien délimitées, bilatérales, et très voisines de la base du crâne.

La visualisation l'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est
excellent une visualisation l'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est
excellent une visualisation l'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est

CONCLUSION :

L'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est
excellent une visualisation l'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est
excellent une visualisation l'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est

L'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est
excellent une visualisation l'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est
excellent une visualisation l'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est

L'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est
excellent une visualisation l'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est
excellent une visualisation l'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est

L'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est
excellent une visualisation l'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est
excellent une visualisation l'articulation du quadrilatère et des muscles du cou, de la nuque, du thorax, du bassin, est

DIEUDONNE M. CALVAT-VIGORON

/

01/01/1967 14 00

01/01/1967 14 00

01/01 14 00

01/01 14 00

01/01 14 00

diabetic patient?

smoker?

able to make his/her
own medical decisions?

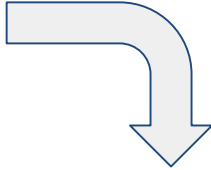
Related to concept extraction,
but implies a
decision/prediction at
document level.

Complex phenotyping: text to database

MAMMOGRAPHY:

There is a **1.8 cm round** **mass** with a circumscribed margin in the **left breast** in the **anterior depth central** to the nipple. There also is a **1.4 cm oval** **mass** with an obscured margin in the **left breast** in the **anterior depth of the inferior region**.

ASSESSMENT: **BI-RADS Category 3**



SCORE	LATERALITY
BI-RADS 3	LEFT
BI-RADS 1	RIGHT

LESION	LATERALITY	LOCALIZATION	SHAPE
mass	LEFT	QSI	1.8
mass	LEFT	QI	1.4

ICD-10 classification

Dr BERNARD CLERC
26 RUE DE LA COMBE
74540 ALBY SUR CHERAN

PÔLE FEMME-MERE-
ENFANT
SERVICE DE PEDIATRIE

Chefs de Service
Dr DIDIER-WRIGHT
Catherine
N° ADEL: 741027002
Dr GLASTRE Catherine
N° ADEL: 741039708

METZ-TESSY, le: 01/02/2011

Praticiens Hospitaliers
Dr BERND Valérie
N° ADEL: 741039384
Dr DELAVIE Nadège
N° ADEL: 741039308
Dr LABIB Fayez
N° ADEL: 741034003
Dr LEVRAT Virginie
N° ADEL: 741043215
Dr MAURE Véronique
N° ADEL: 741033208
Dr PEYRET Blandine
N° ADEL: 741044302
Dr POLES Catherine
N° ADEL: 741039308

Assistante
Dr JARRASSE Clemence
N° ADEL: 741040510

Allergologue
Dr VERMEULEN
Christophe
N° ADEL: 741037302

Cadres de Santé
Service de Néonatalogie -
Soins Intensifs
Tél : 04 50 63 68 04

Service des Nourissons
Tél : 04 50 63 66 67

Service des Grands Enfants
Tél : 04 50 63 66 55

Hôpital de Jour
Tél : 04 50 63 66 57

Secrétariat
Tél : 04 50 63 63 24
Fax : 04 50 63 60 98
pediatrie.secr@ch-anneey.fr

Références : CGL/CGL

Objet : **Compte-rendu d'hospitalisation (Du 30/12/2010 au 08/01/2011)**

Le petit **NOBILI Nello**, né le 17/10/2010 (2 mois), a été hospitalisé le 30/12/2010 dans le service des **NOURRISSONS** pour fièvre évoluant depuis 4 jours avec perte d'appétit et, depuis le jour même, constatation d'éruption maculeuse.

Observation

Anamnèse

Début de la fièvre le 26/12 (38°8 à domicile ; 38° aux urgences pédiatriques après prise de paracétamol) bien tolérée, sans point d'appel infectieux clinique autre qu'un encombrement nasal modéré. Appétit discrètement altéré (5 tétées au lieu de 6 dans la journée).

Biologie faite le 26/12 aux urgences pédiatriques :

- NFP normale
- CRP négative
- cytolysé à 4N (TGO = 121 UI/l ; TGP = 145 UI/l)

Retour à domicile avec conseil de reconsulter en cas de nouveau symptôme.

Apyrexie le 27/12.

Reprise d'un fébricule à 38° à partir du 28/12 et, à partir du 28/12 également, toussé, est encombré et dort mal.

Ne prend plus que des biberons de 30 à 60 ml depuis 4 jours lorsqu'il vient reconsulter le 30/12.

Antécédents

Antécédents personnels :

Né à 34SA + 6 jours . PN = 2690 g .

Examen clinique

Le 30/12 :

- température = 38°
- poids = 4450 g alors qu'il avait été pesé à 4650 g le 26/12 mais ce poids était sans doute surestimé

ICD-10 classification

Dr BERNARD CLERC
26 RUE DE LA COMBE
74540 ALBY SUR CHERAN

PÔLE FEMME-MÈRE-
ENFANT
SERVICE DE PÉDIATRIE

Chefs de Service
Dr DIDIER-WRIGHT
Catherine
N° ADEL: 741027002
Dr GLASTRE Catherine
N° ADEL: 741039708

Praticiens Hospitaliers
Dr BENO Valérie
N° ADEL: 741030384
Dr DELAVIE Nadège
N° ADEL: 741030308
Dr LABIB Fayez
N° ADEL: 741034003
Dr LEVRAT Virginie
N° ADEL: 741040215
Dr MAURE Véronique
N° ADEL: 741030208
Dr PEYRET Blandine
N° ADEL: 741044302
Dr POLES Catherine
N° ADEL: 741030308

Assistante
Dr JARRASSE Clemence
N° ADEL: 741040510

Allergologue
Dr VERMEULEN
Christophe
N° ADEL: 741037302

Cadres de Santé
Service de Néonatalogie -
Soins Intensifs
Tél : 04 50 63 68 04

Service des Nourissons
Tél : 04 50 63 66 67

Service des Grands Enfants
Tél : 04 50 63 66 55

Hôpital de Jour
Tél : 04 50 63 66 57

Secrétariat
Tél : 04 50 63 63 24
Fax : 04 50 63 60 98
pediatrie.secr@ch-annecy.fr

METZ-TESSY, le 01/02/2011



Classification, ICD-10 codes

- ▶ VII Diseases of the eye and adnexa
- ▶ VIII Diseases of the ear and mastoid process
- ▶ IX Diseases of the circulatory system
- ▼ X Diseases of the respiratory system
 - ▶ J00-J06 Acute upper respiratory infections
 - ▼ J09-J18 Influenza and pneumonia
 - J09 Influenza due to identified zoonotic or pandemic influenza virus
 - ▶ J10 Influenza due to identified seasonal influenza virus
 - ▶ J11 Influenza, virus not identified
 - ▼ J12 Viral pneumonia, not elsewhere classified
 - J12.0 Adenoviral pneumonia
 - J12.1 Respiratory syncytial virus pneumonia
 - J12.2 Parainfluenza virus pneumonia
 - J12.3 Human metapneumovirus pneumonia
 - J12.8 Other viral pneumonia
 - J12.9 Viral pneumonia, unspecified



ICD-10 classification

Dr BERNARD CLERC
26 RUE DE LA COMBE
74540 ALBY SUR CHERAN

PÔLE FEMME-MÈRE-
ENFANT
SERVICE DE PÉDIATRIE

Chefs de Service
Dr DIDIER-WRIGHT
Catherine
N° ADEL: 741027002
Dr GLASTRE Catherine
N° ADEL: 741039708

Praticiens Hospitaliers
Dr BERND Valérie
N° ADEL: 741039384
Dr DELAVIE Nadège
N° ADEL: 741039308
Dr LABIB Fayez
N° ADEL: 741034003
Dr LEVRAT Virginie
N° ADEL: 741043215
Dr MAURE Véronique
N° ADEL: 741033208
Dr PEYRET Blandine
N° ADEL: 741044302
Dr POLES Catherine
N° ADEL: 741033036

Assistante
Dr JARRASSE Clemence
N° ADEL: 741040510

Allergologue
Dr VERMEULEN
Christophe
N° ADEL: 741037302

Cadres de Santé
Service de Néonatalogie -
Soins Intensifs
Tél : 04 50 63 68 04

Service des Nourissons
Tél : 04 50 63 66 67

Service des Grands Enfants
Tél : 04 50 63 66 55

Hôpital de Jour
Tél : 04 50 63 66 57

Secrétariat
Tél : 04 50 63 63 24
Fax : 04 50 63 60 98
pediatrie.secr@ch-annecy.fr

METZ-TESSY, le 01/02/2011



Classification, ICD-10 codes

- Big but noisy data
- 14k+ classes
 - Often several codes per visit
 - Some classes are rare or never used
 - Hierarchical classification
- Heterogeneous data
 - Textual data
 - Structured data

Références : CGL/CGL

Objet : **Compte-rendu d'hospitalisation (Du 30/12/2010 au 08/01/2011)**

Le petit **NOBILI Nello**, né le 17/10/2010 (2 mois), a été hospitalisé le 30/12/2010 dans le service des **NOURRISSONS** pour fièvre évoluant depuis 4 jours avec perte d'appétit et, depuis le jour même, constatation d'éruption maculeuse.

Observation

Anamnèse

Début de la fièvre le 26/12 (38°8 à domicile ; 38° aux urgences pédiatriques après prise de paracétamol) bien tolérée, sans point d'appel infectieux clinique autre qu'un encombrement nasal modéré. Appétit discrètement altéré (5 tétées au lieu de 6 dans la journée).

Biologie faite le 26/12 aux urgences pédiatriques :

- NFP normale
- CRP négative
- cytolysé à 4N (TGO = 121 UI/l ; TGP = 145 UI/l)

Retour à domicile avec conseil de reconsulter en cas de nouveau symptôme.

Apyrexie le 27/12.

Reprise d'un fébricule à 38° à partir du 28/12 et, à partir du 28/12 également, toussé, est encombré et dort mal.

Ne prend plus que des biberons de 30 à 60 ml depuis 4 jours lorsqu'il vient reconsulter le 30/12.

Antécédents

Antécédents personnels :

Né à 34SA + 6 jours . PN = 2690 g .

Examen clinique

Le 30/12 :

- température = 38°
- poids = 4450 g alors qu'il avait été pesé à 4650 g le 26/12 mais ce poids était sans doute surestimé

Overall

What makes a task difficult?

Difficult to solve by a human being/expert

(cf. inter/intra-annotator agreement, annotation good practices)

multi-class

(e.g: find all medical concepts in a clinical report)

multi-label

(ex. : ICD-10 classification)

Skewed, unbalanced distribution of classes

(e.g.: ill/not ill)

Complex representation

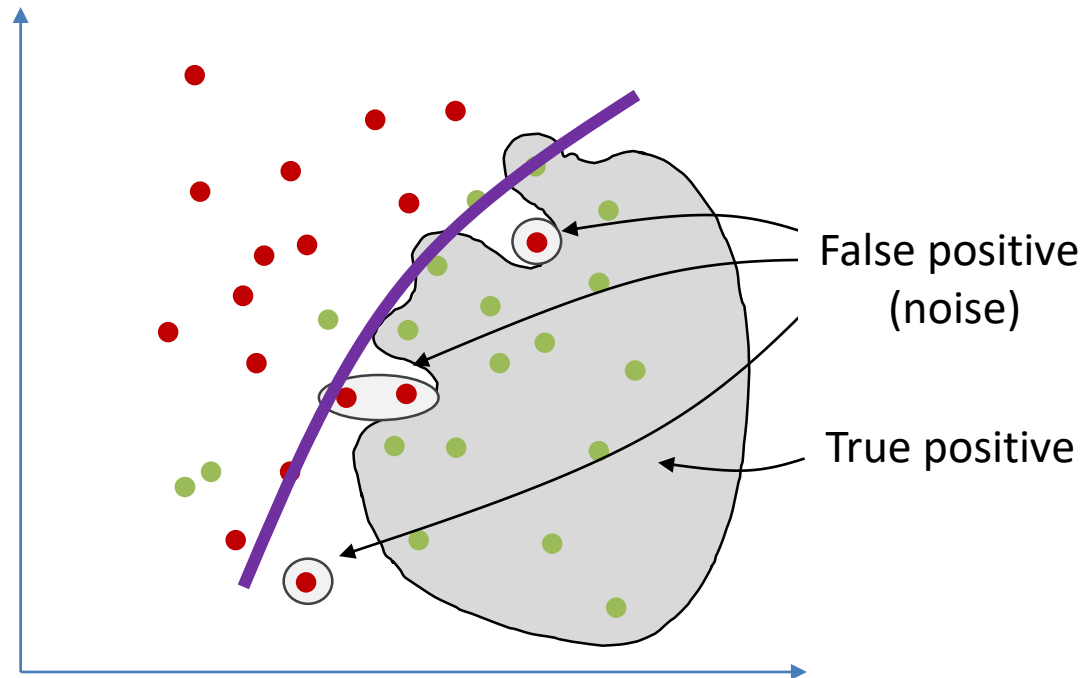
(e.g.: text – ambiguous, implicit, redundant, produced by humans for humans)

Missing and/or noisy data

(e.g.: almost all real-life problems)

Metrics

For the class « GREEN »

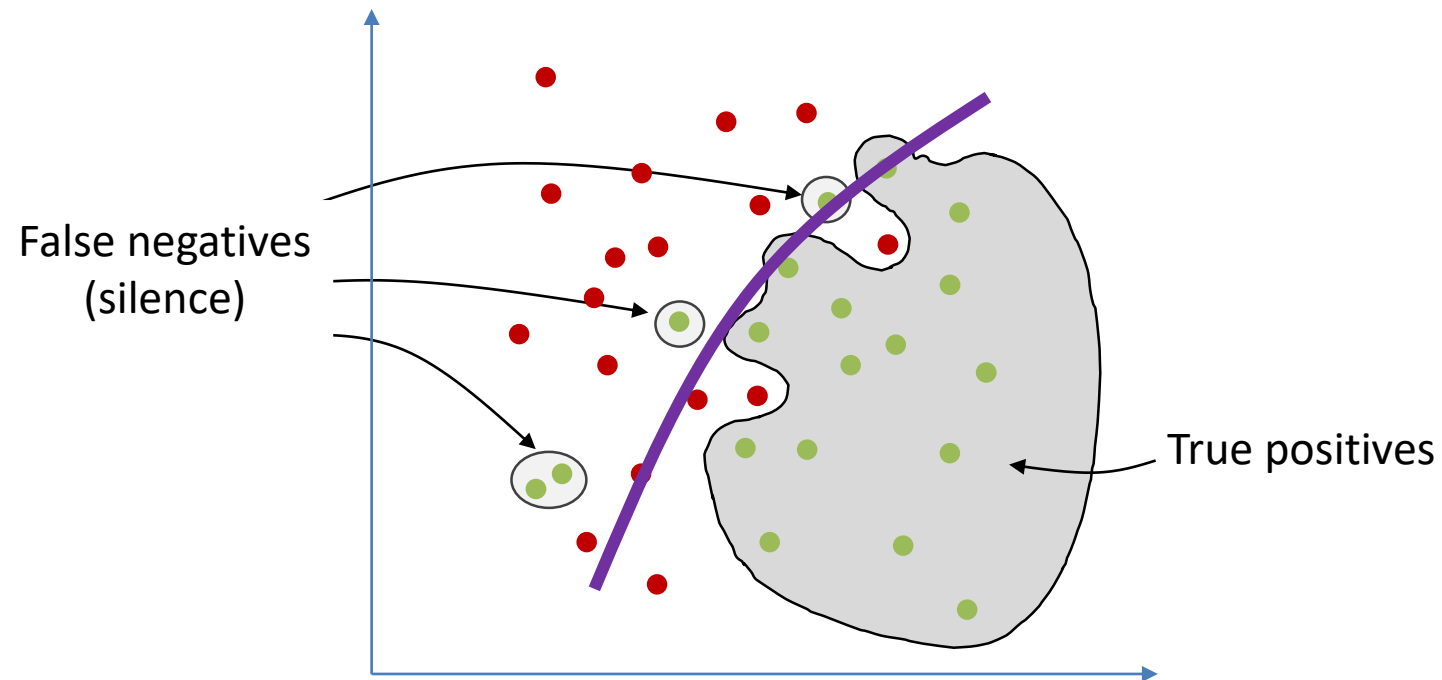


$$\text{Positive predictive value} = \text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}} = 1 - \text{noise}$$

A good prediction rarely predicts that a healthy patient is ill.

Metrics

For the class « GREEN »

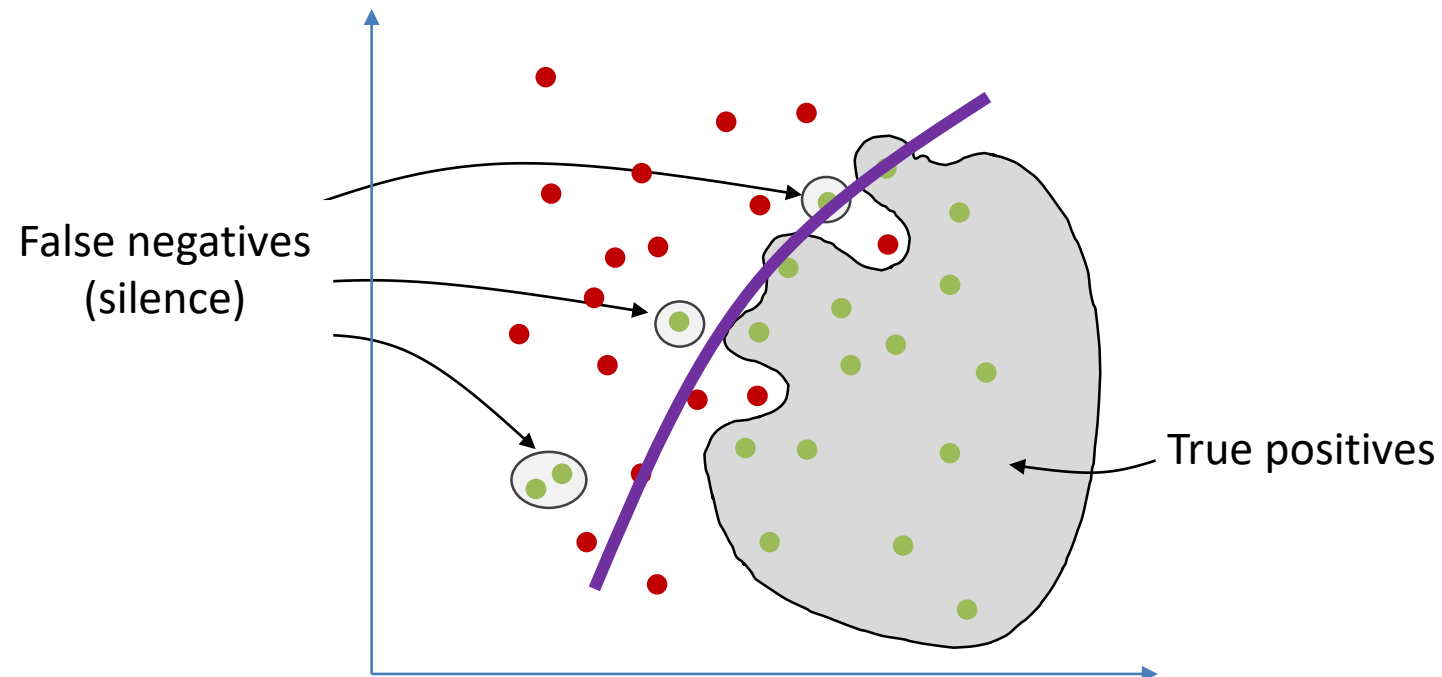


$$\text{Sensitivity} = \text{recall} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}} = 1 - \text{silence}$$

A good prediction rarely predicts that an ill patient is healthy.

Metrics

For the class « GREEN »



$$\text{Specificity} = \frac{\text{True negatives}}{\text{True negatives} + \text{False positives}} = 1 - \text{silence}$$

A good prediction rarely predicts that a healthy patient is ill.

Metrics

- **Accuracy** = rate of good predictions, but very prone to misinterpretation!

		Predictions	
		Healthy	Ill
Gold standard	Healthy	19064	89
	Ill	110	231

- Accuracy = 0.99
- Recall (ill) = 0.68
- Precision (ill) = 0,72

Machine learning vs the real world

- Machine learning = it's all about prediction & find answers to questions
- But what about
 - Causality
 - Feedback effect
 - Interpretability
 - Reliability
 - Replicability
 - Robustness
 - Adaptability
 - Ethics
 - ?