SIMCA in Crime Scene Investigations

Dr. Laura Ortiz Herrero and Dr. Luis Bartolomé Moro, University of the Basque Country (UPV/EHU)

Umetrics Webinar May 2021

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea
ERTZAINTZA
sgiker
Ikerkuntzarako Zerbitzu Orokorrak
Servicios Generales de Investigación
POLICIA JUDICIARIA
FCTUC FACULDADE DE CIÊNCIAS E TECNOLOGIA UNIVERSIDADE DE COIMBRA

Outline

Introduction

- Importance of determining the time frame at the crime scene.
- Potential of chemometrics and current status of its implementation to forensic dating.
- Objective.

Methodology

- Proposed chemometric based dating methodology.

Case study

- Questioned documents field.

Conclusions

- General conclusions drawn from the thesis work.

When did an event **occur**?

When was a trace **left** on the crime scene?

How old is an evidence?

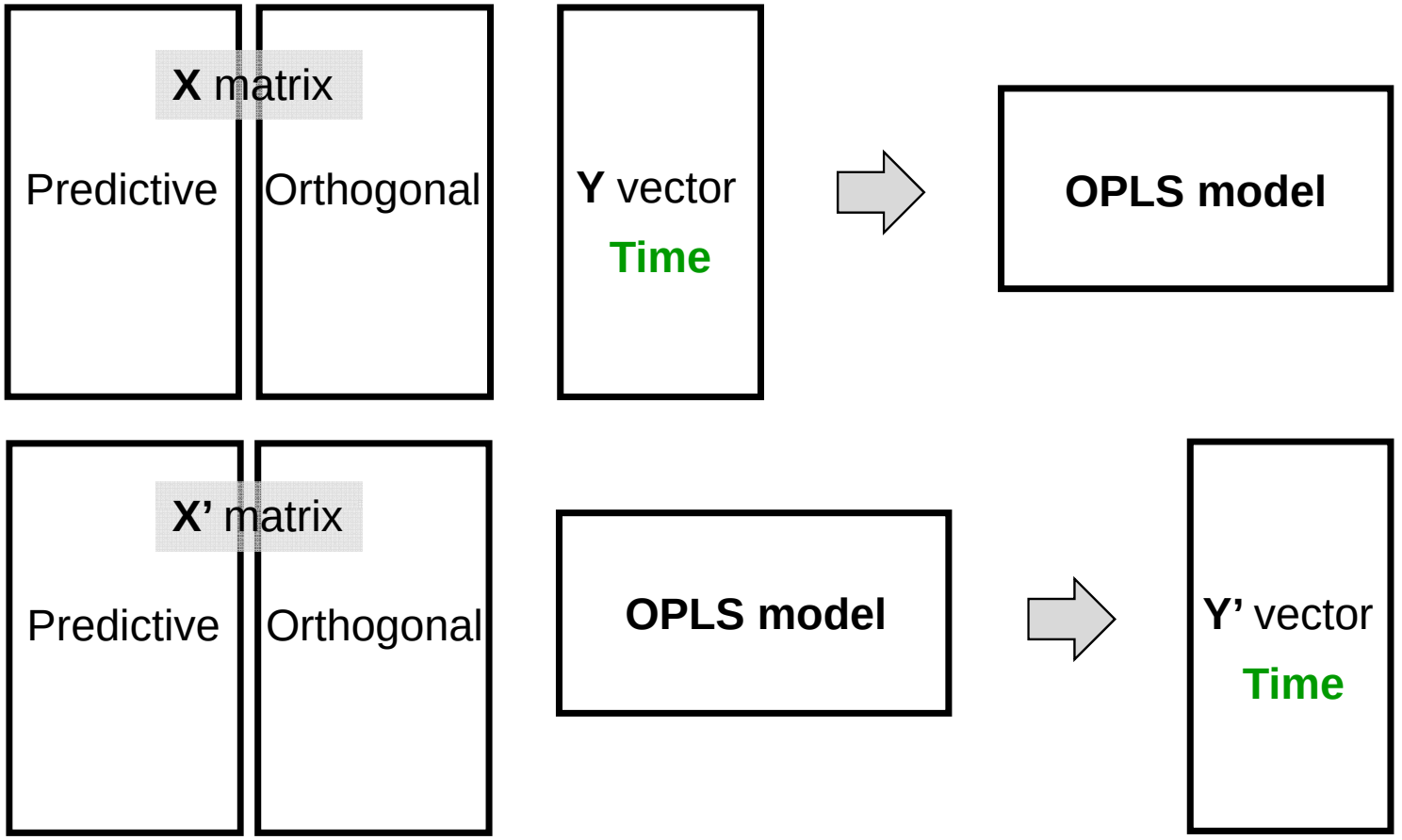
In what order did events **occur**?

Time has been largely **unexplored** due to the complexity of the **overall challenge.**

combined with **chemometrics**

- ✓ Improves **handling** and **interpretation** of high dimensional data.
- ✓ Provides additional **information**.
- ✓ Achieves objective and meaningful **results**.
- ✓ **Fast** in the time domain.
- ✓ Reduces human **error**.

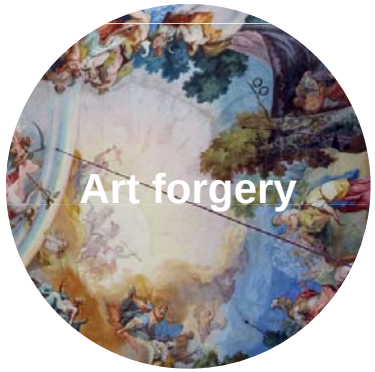
Partial least squares regression (PLSR) and its **extension OPLSR**.



Chemometrics has **recently** been **introduced** to forensic dating.

- × **Potential applicability** poorly understood and underexploited.
- × Lack of **statistical expertise** for its optimal application.
- × Limited implementation of **standardised** chemometric-based **methodologies** in forensic laboratories.

The **aim** of this thesis has been to **highlight the usefulness of multivariate regression methods** in the **forensic field** through the **development and validation of dating methodologies** in which non-destructive and micro-destructive techniques have been applied together with the (O)PLSR method.



Journal of Analytical and Applied Pyrolysis 131 (2018) 9-16

Contents lists available at ScienceDirect

Journal of Analytical and Applied Pyrolysis

journal homepage: www.elsevier.com/locate/jaap

Direct and indirect approaches based on paper analysis by Py-GC/MS for estimating the age of documents

L. Ortiz-Herrero[a,*], M.E. Blanco[a], C. García-Ruiz[b,c], L. Bartolomé[d]

Microchemical Journal 140 (2018) 158-166

Contents lists available at ScienceDirect

Microchemical Journal

journal homepage: www.elsevier.com/locate/microc

DATUVINK pilot study: A potential non-invasive methodology for dating ballpoint pen inks using multivariate chemometrics based on their UV–vis-NIR reflectance spectra

L. Ortiz-Herrero[a,*], L. Bartolomé[b], I. Durán[a], I. Velasco[a], M.L. Alonso[a], M.I. Maguregui[c], M. Ezcurra[d]

Chemometrics and Intelligent Laboratory Systems 207 (2020) 104187

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemometrics

A novel, non-invasive, multi-purpose and comprehensive method to date inks in real handwritten documents based on the monitoring of the dye ageing processes

L. Ortiz-Herrero [a,*], A.C. de Almeida Assis [b], L. Bartolomé [c], M.L. Alonso [a], M.I. Maguregui [d], R.M. Alonso [a], J.S. Seixas de Melo [e]

Talanta 205 (2019) 120114

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

OPLS multivariate regression of FTIR-ATR spectra of acrylic paints for age estimation in contemporary artworks

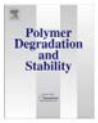
L. Ortiz-Herrero[a,*], I. Cardaba[b], S. Setien[a], L. Bartolomé[c], M.L. Alonso[a], M.I. Maguregui[b]

Polymer Degradation and Stability 179 (2020) 109263

Contents lists available at ScienceDirect

Polymer Degradation and Stability

journal homepage: www.elsevier.com/locate/polydegstab



Extension study of a statistical age prediction model for acrylic paints

L. Ortiz-Herrero [a,*], I. Cardaba [b], L. Bartolomé [c], M.L. Alonso [a], M.I. Maguregui [b]

Trends in Analytical Chemistry 141 (2021) 116278

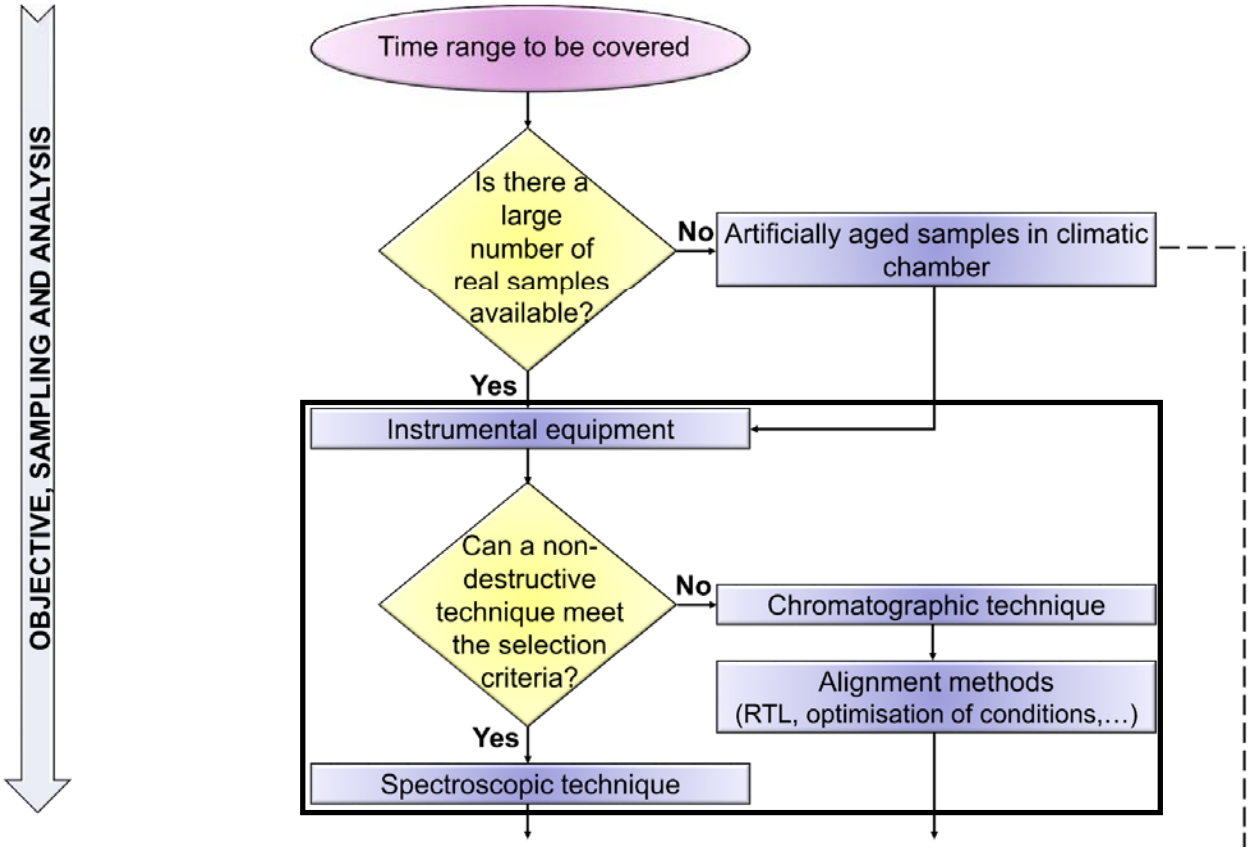
Contents lists available at ScienceDirect

Trends in Analytical Chemistry

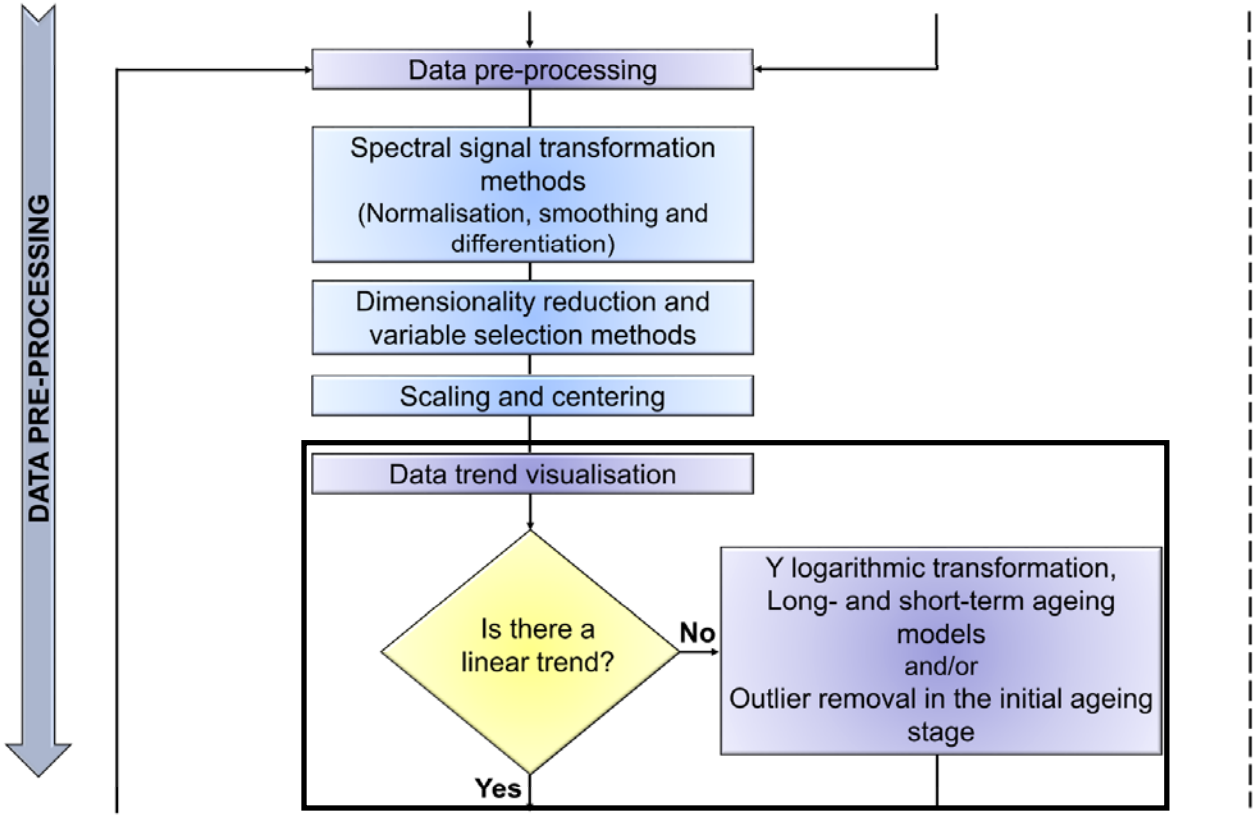
journal homepage: www.elsevier.com/locate/trac

Multivariate (O)PLS regression methods in forensic dating

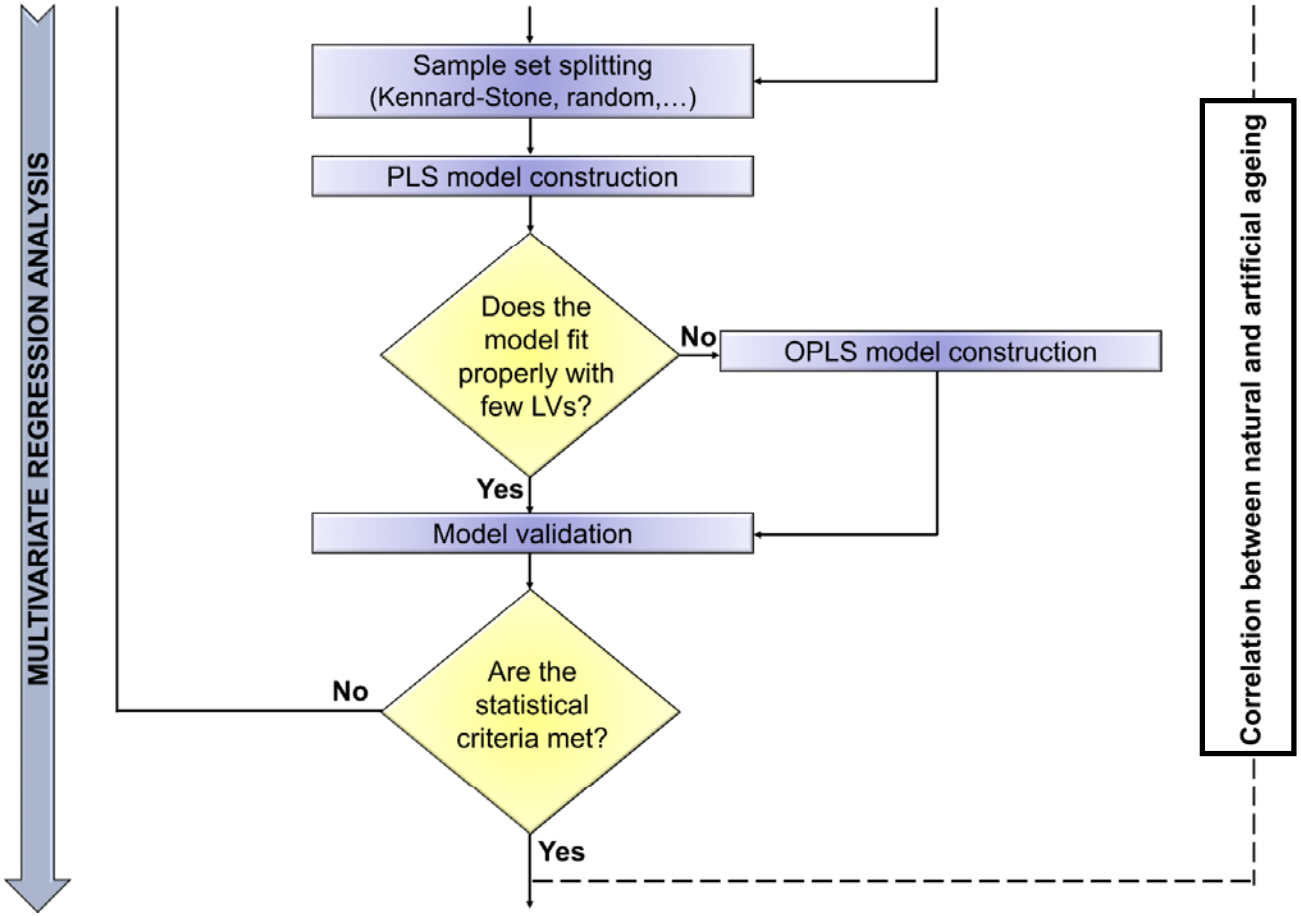
L. Ortiz-Herrero [a,*], M.I. Maguregui [b], L. Bartolomé [c]

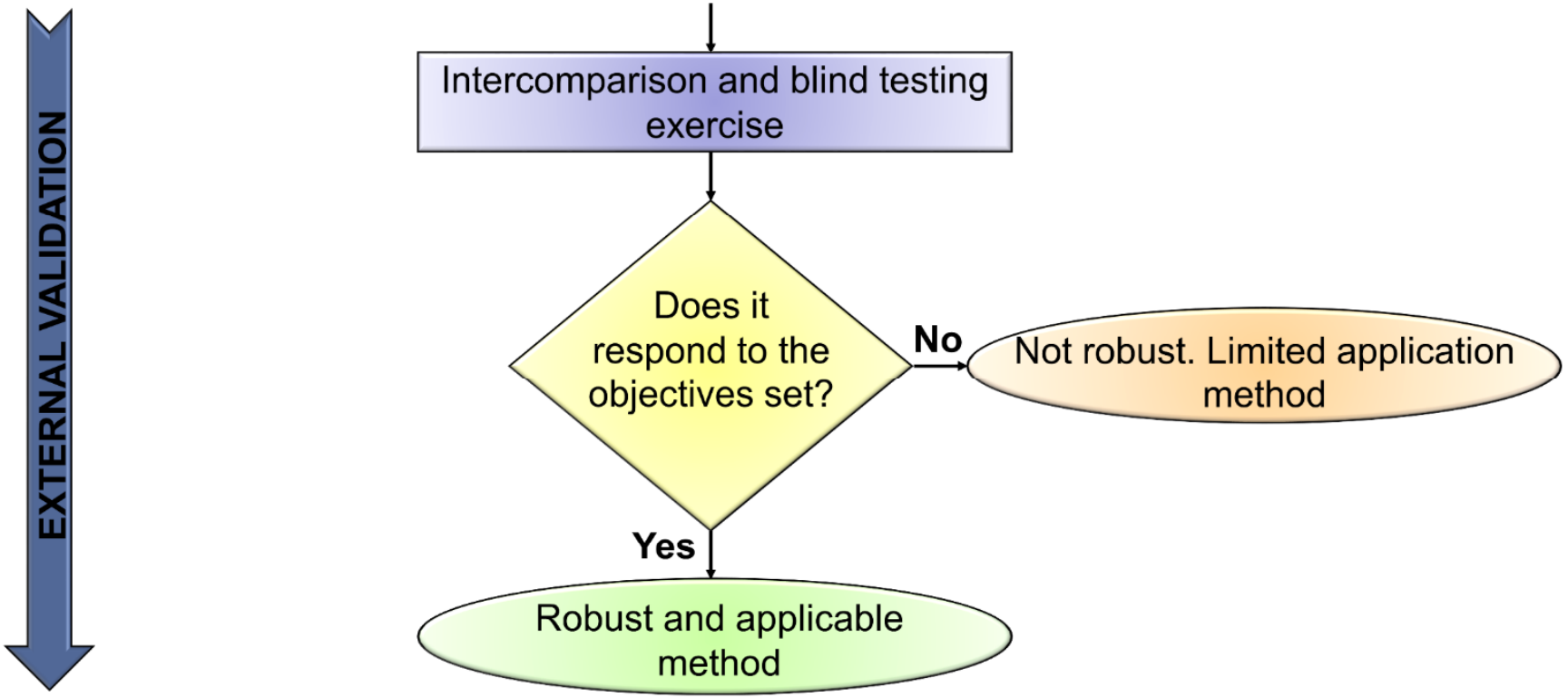

OBJECTIVE, SAMPLING AND ANALYSIS
Time range to be covered
Is there a large number of real samples available?
No
Artificially aged samples in climatic chamber
Yes
Instrumental equipment
Can a non-destructive technique meet the selection criteria?
No
Chromatographic technique
Alignment methods (RTL, optimisation of conditions,…)
Yes
Spectroscopic technique

SIMCA 15.0.2 Umetrics® software



SIMCA 15.0.2 Umetrics® software



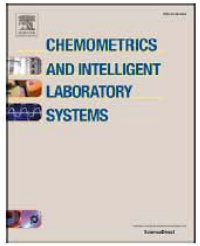

EXTERNAL VALIDATION
Intercomparison and blind testing exercise
Does it respond to the objectives set?
No
Not robust. Limited application method
Yes
Robust and applicable method

Chemometrics and Intelligent Laboratory Systems 207 (2020) 104187

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemometrics

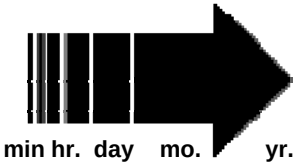


A novel, non-invasive, multi-purpose and comprehensive method to date inks in real handwritten documents based on the monitoring of the dye ageing processes

L. Ortiz-Herrero [a,*], A.C. de Almeida Assis [b], L. Bartolomé [c], M.L. Alonso [a], M.I. Maguregui [d], R.M. Alonso [a], J.S. Seixas de Melo [e]

[a] Analytical Chemistry Department, Faculty of Science and Technology, University of the Basque Country (UPV/EHU), Barrio Sarriena S/N, 48940, Leioa, Bizkaia, Spain
[b] Scientific Police Laboratory, Judiciary Police, Rua Gomes Freire, 174, 1169-007, Lisbon, Portugal
[c] Advances Research Facilities (SGIker), Martina Casiano Technology Platform, University of the Basque Country (UPV/EHU), Barrio Sarriena S/N, 48940, Leioa, Bizkaia, Spain
[d] Painting Department, Faculty of Fine Arts, University of the Basque Country (UPV/EHU), Barrio Sarriena S/N, 48940, Leioa, Bizkaia, Spain
[e] Coimbra Chemistry Centre, Department of Chemistry, University of Coimbra, Rua Larga, 3004-535, Coimbra, Portugal

Requirements to be met by the **methodologies** to be developed:



Drawbacks encountered in the development of these methodologies:

Objective

- ✓ Method for **dating handwritten documents**.
- ✓ **Accurate** prediction of the ink stroke.
- ✓ Applied to the **largest number** of pen **inks** and to a **wide time range**.
- ✓ **Non-invasive** measurement by Vis-MSP.
- ✓ **Reliable** results with 95% confidence.
- ✓ **Minimal influence** of paper **support** and preservation **conditions**.

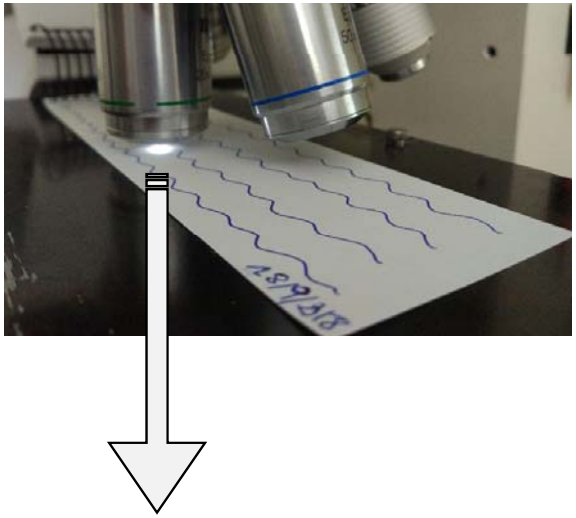
Ink sampling

Pen brand	Pen type	No. of samples	Temporary range of application (mo.)
Bic® Cristal Medium France	Ballpoint	25	25
Bic® Cristal Medium USA	Ballpoint	30	27
Staedtler® Stick 430 M	Ballpoint	27	23
Paper Mate® Flex Grip Elite 1.4 M	Ballpoint	20	25
Paper Mate® 0.7mm	Liquid (gel ink)	20	25
Pilot® Super Grip M	Ballpoint	25	25
Pilot® G-1	Liquid (gel ink)	20	25
Faber-Castell® Medium	Ballpoint	28	23
Uni-Ball® Signo Broad	Liquid (gel ink)	23	25
Uni-Ball® Jetstream Sport	Ballpoint	21	25
Inoxcrom® TC Ball Stainless Steel Tip M	Liquid (gel ink)	22	25

11 writing tools from **7 brands**

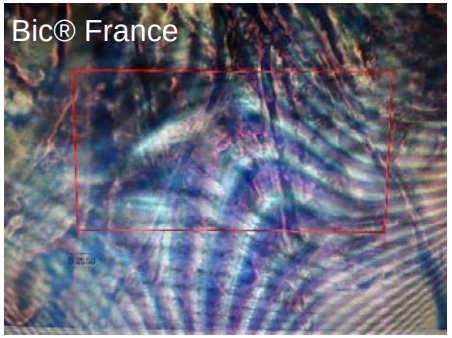
Max. **2.5 natural years**

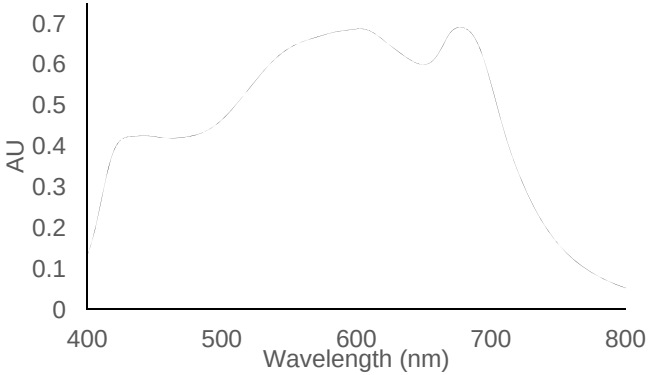
Vis-MSP analysis



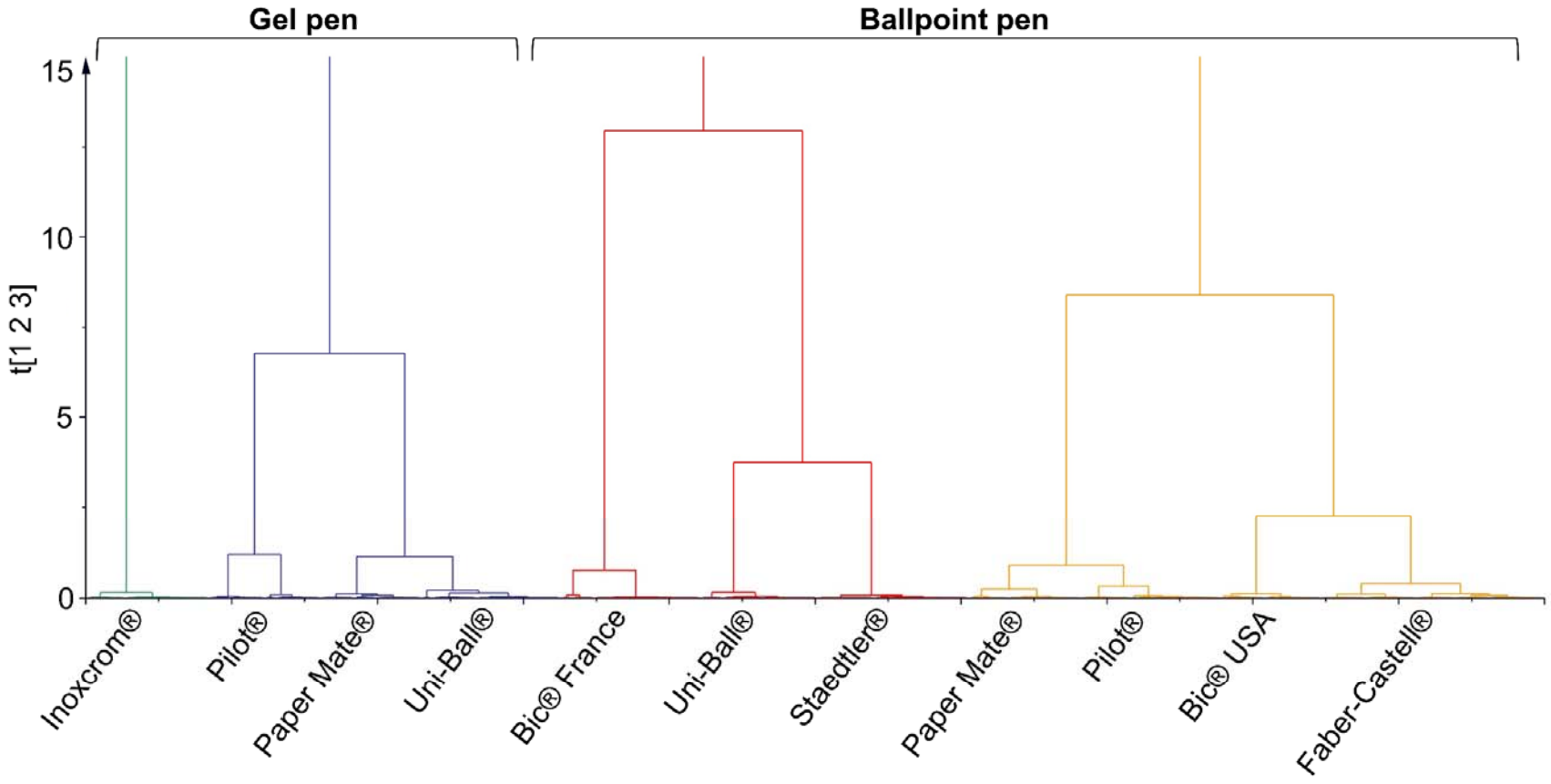
Spectrophotometer J&M TIDAS® MSP-800
Interpolation yes
1 nm resolution
Single scan, 3 accumulations
1 pixel bunching
400-800 nm range, reflectance acquisition mode

Zeiss® Axiotech 100 Microscope
220.0 x 127.0 µm diaphragm
640 x 80 image resolution
20x objective magnification
10x light intensity





Pre-classification by HCA

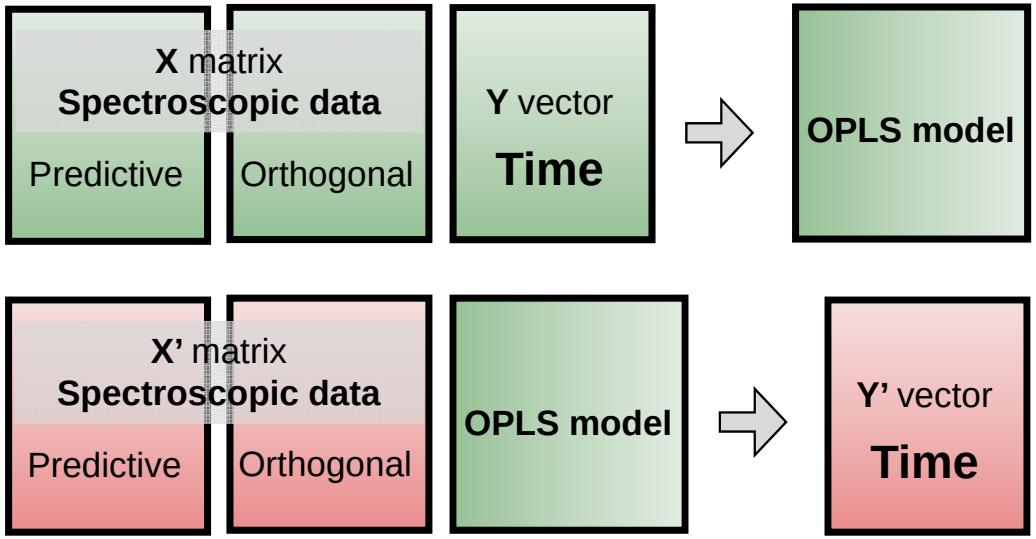


OPLS model building and validation

Ink type	Pen brand
Ballpoint	Bic® France
	Bic® USA
	Faber-Castell®
	Staedtler®
	Pilot®
	Paper Mate®
	Uni-Ball®
Gel	Paper Mate®
	Uni-Ball®
	Pilot®
	Inoxcrom®

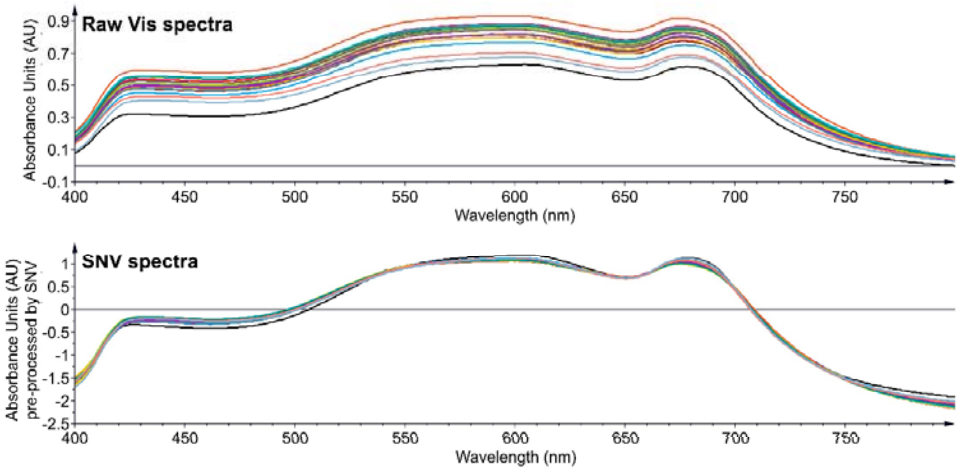
SIMCA V. 15.0.2 Umetrics® software, Stone algorithm, 80% training and 20% test

One OPLS model per writing instrument (n=11)



OPLS model building and validation

Ink type	Pen brand	No. calibration/validation samples	Pre-treatment (scale/filter)
Ballpoint	Bic® France	20/5	UV and SNV
	Bic® USA	23/7	Centring and SNV
	Faber-Castell®	21/7	UV and MSC
	Staedtler®	20/7	UV and 1st derivative
	Pilot®	20/5	UV and SNV
	Paper Mate®	15/5	UV and 1st derivative
	Uni-Ball®	16/5	UV and MSC
Gel	Paper Mate®	15/5	UV and SNV
	Uni-Ball®	18/5	UV and 1st derivative
	Pilot®	15/5	UV and MSC
	Inoxcrom®	17/5	UV and MSC



Remove systematic **non-time-related** variation in X

OPLS model building and validation

Model fitting — Internal predictive testing by cross-validation

Ink type	Pen brand	No. calibration/validation samples	Pre-treatment (scale/filter)	LV	R^2Y	Q^2	RMSEE (days)	RMSECV (days)
Ballpoint	Bic® France	20/5	UV and SNV	4	0.87	0.74	1.2	1.3
	Bic® USA	23/7	Centring and SNV	4	0.86	0.74	1.2	1.3
	Faber-Castell®	21/7	UV and MSC	2	0.71	0.52	1.4	1.4
	Staedtler®	20/7	UV and 1st derivative	6	0.95	0.83	1.2	1.3
	Pilot®	20/5	UV and SNV	1	0.65	0.55	1.4	1.4
	Paper Mate®	15/5	UV and 1st derivative	3	0.78	0.53	1.4	1.5
	Uni-Ball®	16/5	UV and MSC	3	0.82	0.63	1.3	1.3
Gel	Paper Mate®	15/5	UV and SNV	3	0.70	0.41	1.2	1.3
	Uni-Ball®	18/5	UV and 1st derivative	2	0.58	0.41	1.5	1.5
	Pilot®	15/5	UV and MSC	1	0.60	0.45	1.2	1.3
	Inoxcrom®	17/5	UV and MSC	3	0.50	0.31	1.4	1.4

OPLS model building and validation

$$E\% = \left(\frac{(y_{real} - y_{predicted})}{y_{real}}\right) x\ 100$$

External predictive testing

Ink type	Pen brand	No. calibration/validation samples	Pre-treatment (scale/filter)	LV	R^2Y	Q^2	RMSEE (days)	RMSECV (days)	RMSEP (days)	Accuracy error (%)
Ballpoint	Bic® France	20/5	UV and SNV	4	0.87	0.74	1.2	1.3	1.1	10
	Bic® USA	23/7	Centring and SNV	4	0.86	0.74	1.2	1.3	1.1	10
	Faber-Castell®	21/7	UV and MSC	2	0.71	0.52	1.4	1.4	1.2	15
	Staedtler®	20/7	UV and 1st derivative	6	0.95	0.83	1.2	1.3	1.2	15
	Pilot®	20/5	UV and SNV	1	0.65	0.55	1.4	1.4	1.2	12
	Paper Mate®	15/5	UV and 1st derivative	3	0.78	0.53	1.4	1.5	1.3	18
	Uni-Ball®	16/5	UV and MSC	3	0.82	0.63	1.3	1.3	1.3	22
Gel	Paper Mate®	15/5	UV and SNV	3	0.70	0.41	1.2	1.3	1.3	13
	Uni-Ball®	18/5	UV and 1st derivative	2	0.58	0.41	1.5	1.5	1.4	25
	Pilot®	15/5	UV and MSC	1	0.60	0.45	1.2	1.3	1.3	16
	Inoxcrom®	17/5	UV and MSC	3	0.50	0.31	1.4	1.4	1.2	15

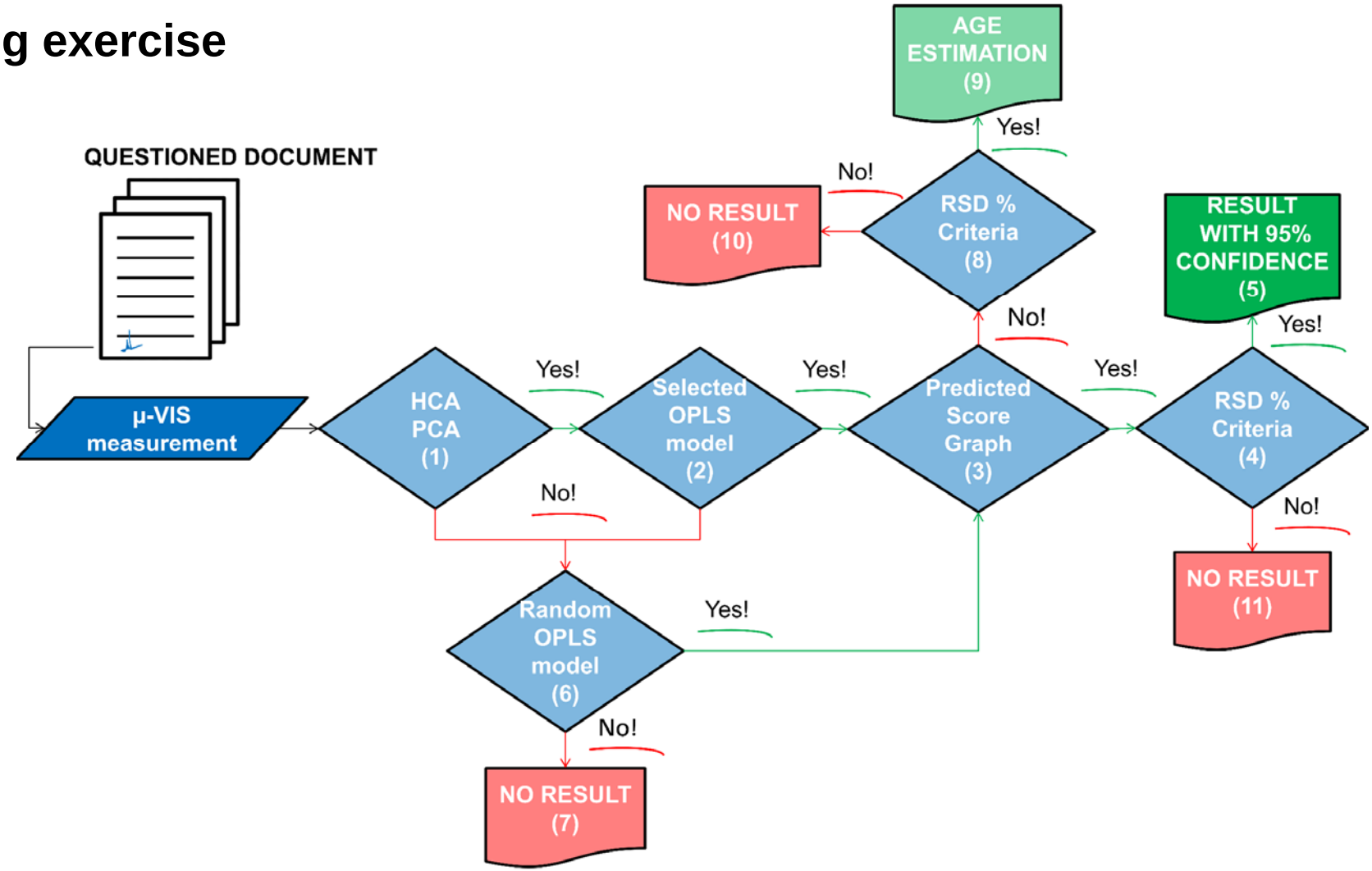
Blind testing exercise

SDG-Q7
F12018

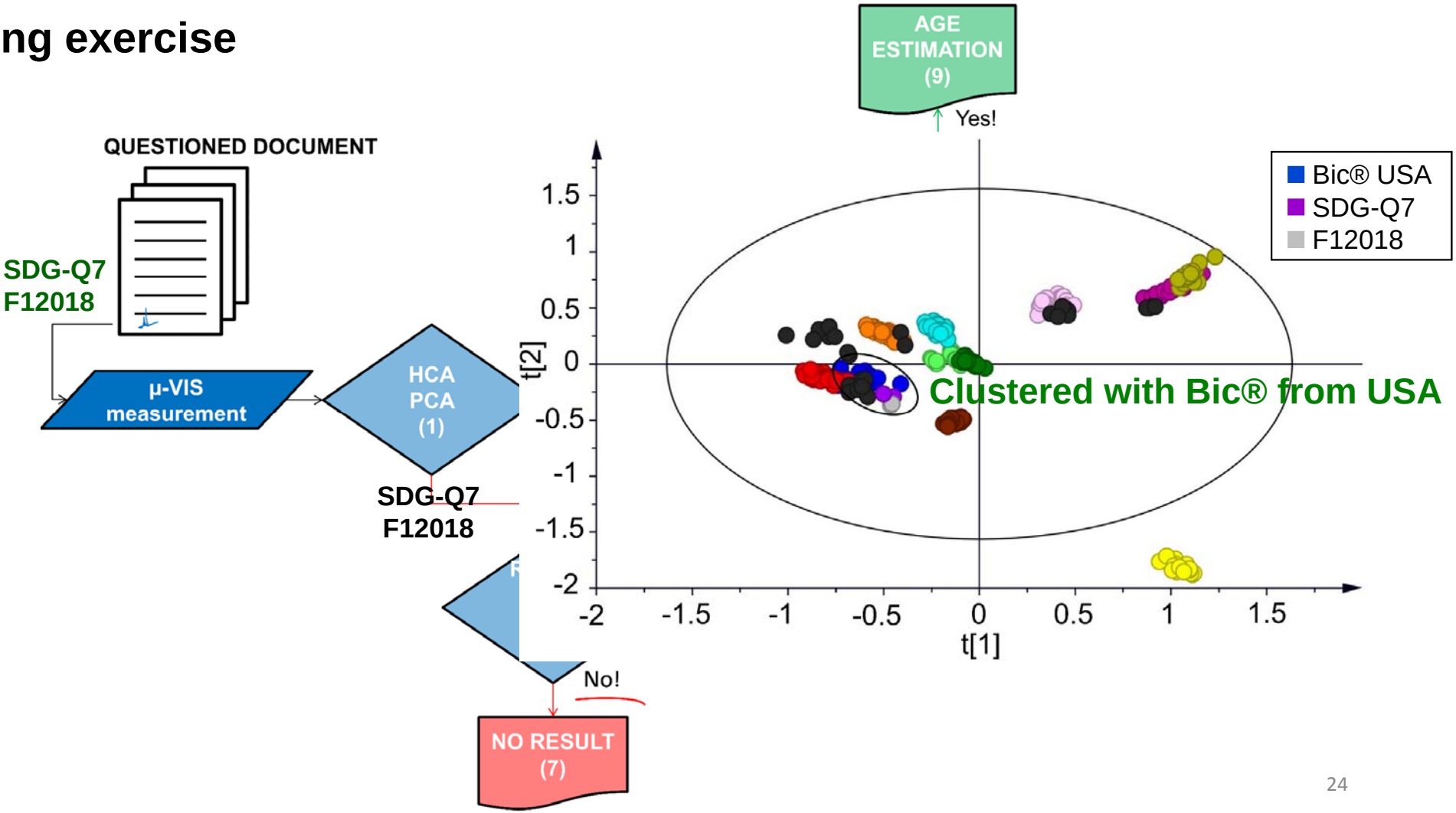
F22018
F32018
F42018

SDG-Q6
SDG-Q8

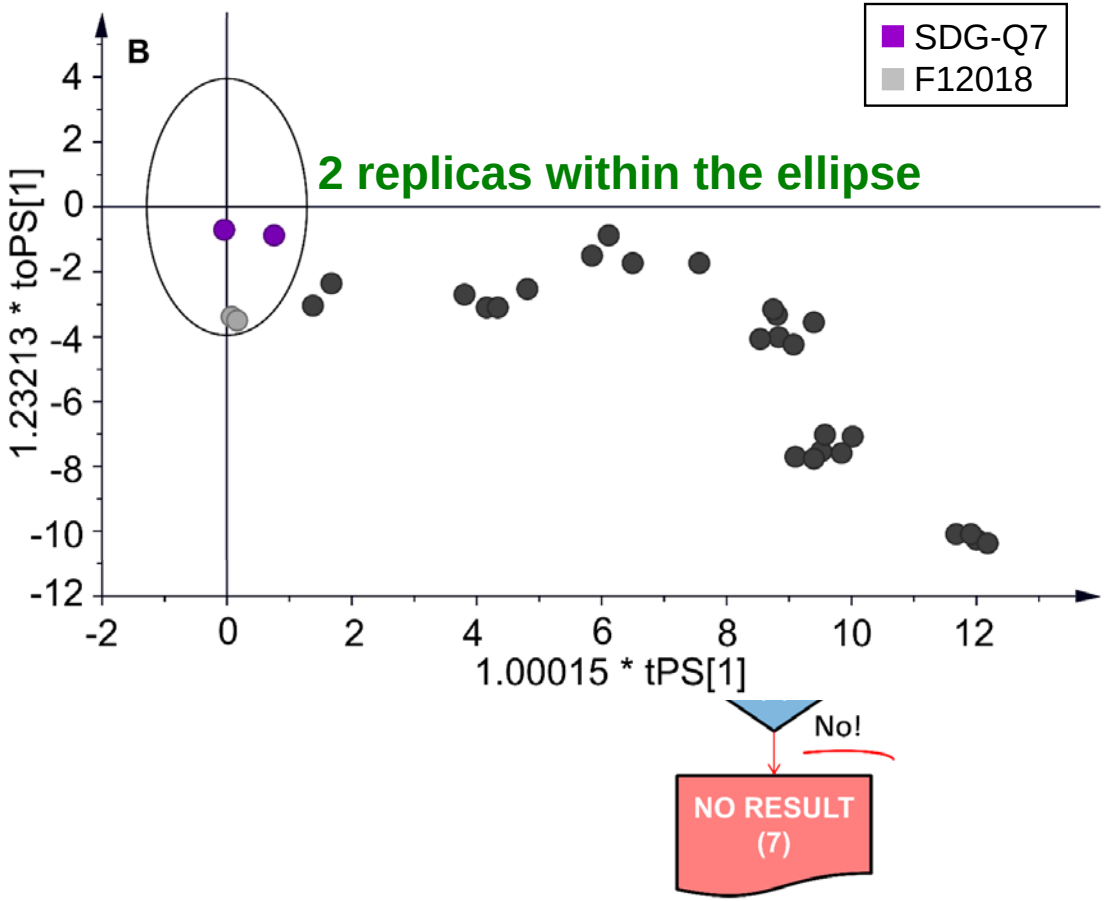
F12019
F22019
F12017
F32017
F42017
F52017

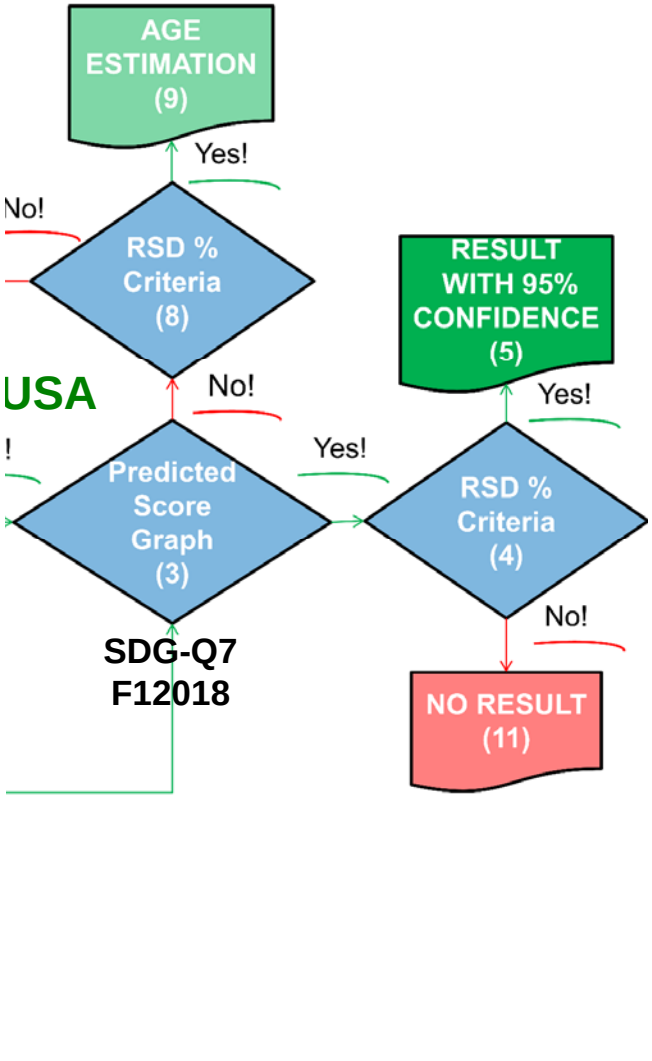


Blind testing exercise

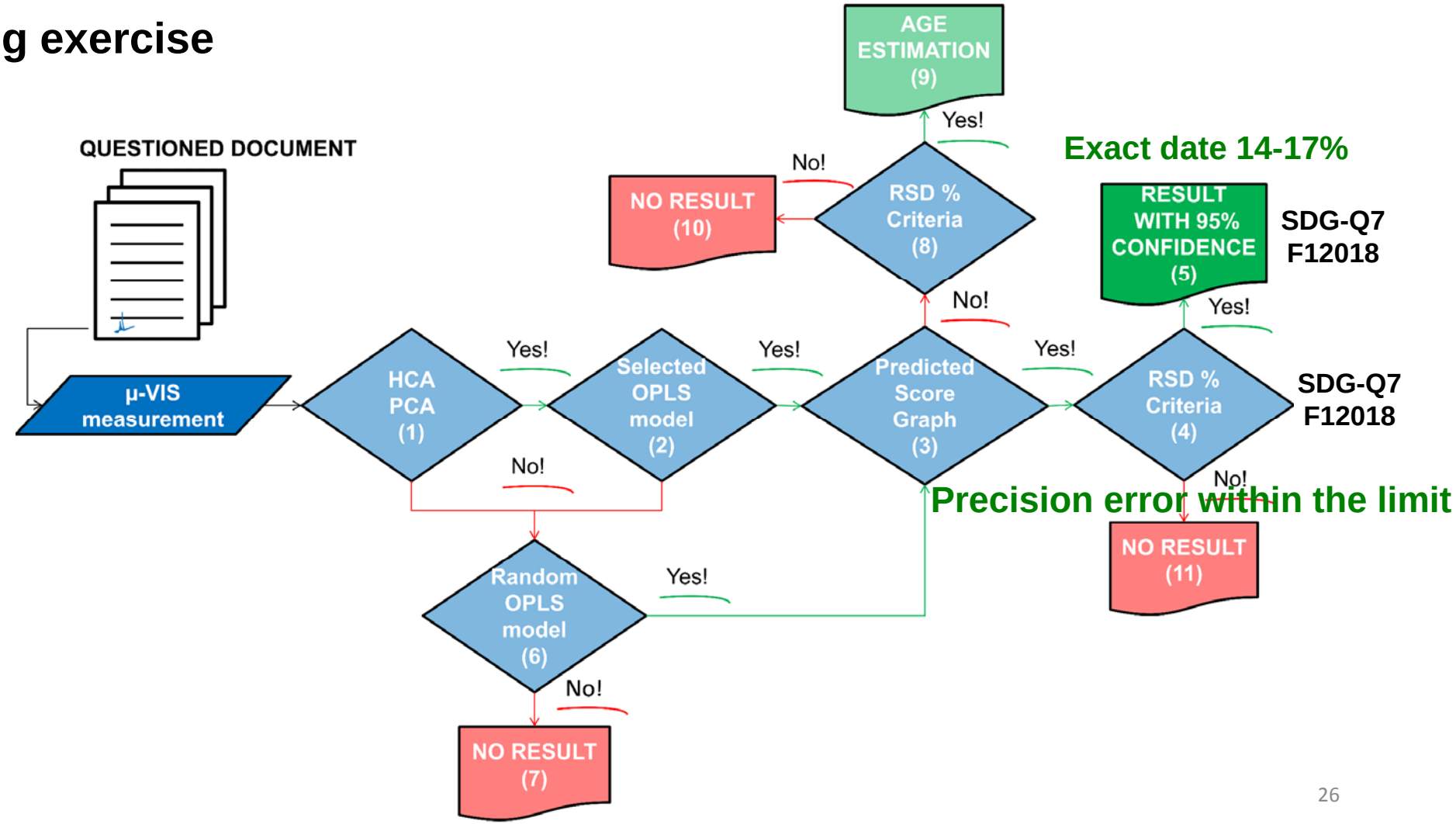


Blind testing exercise

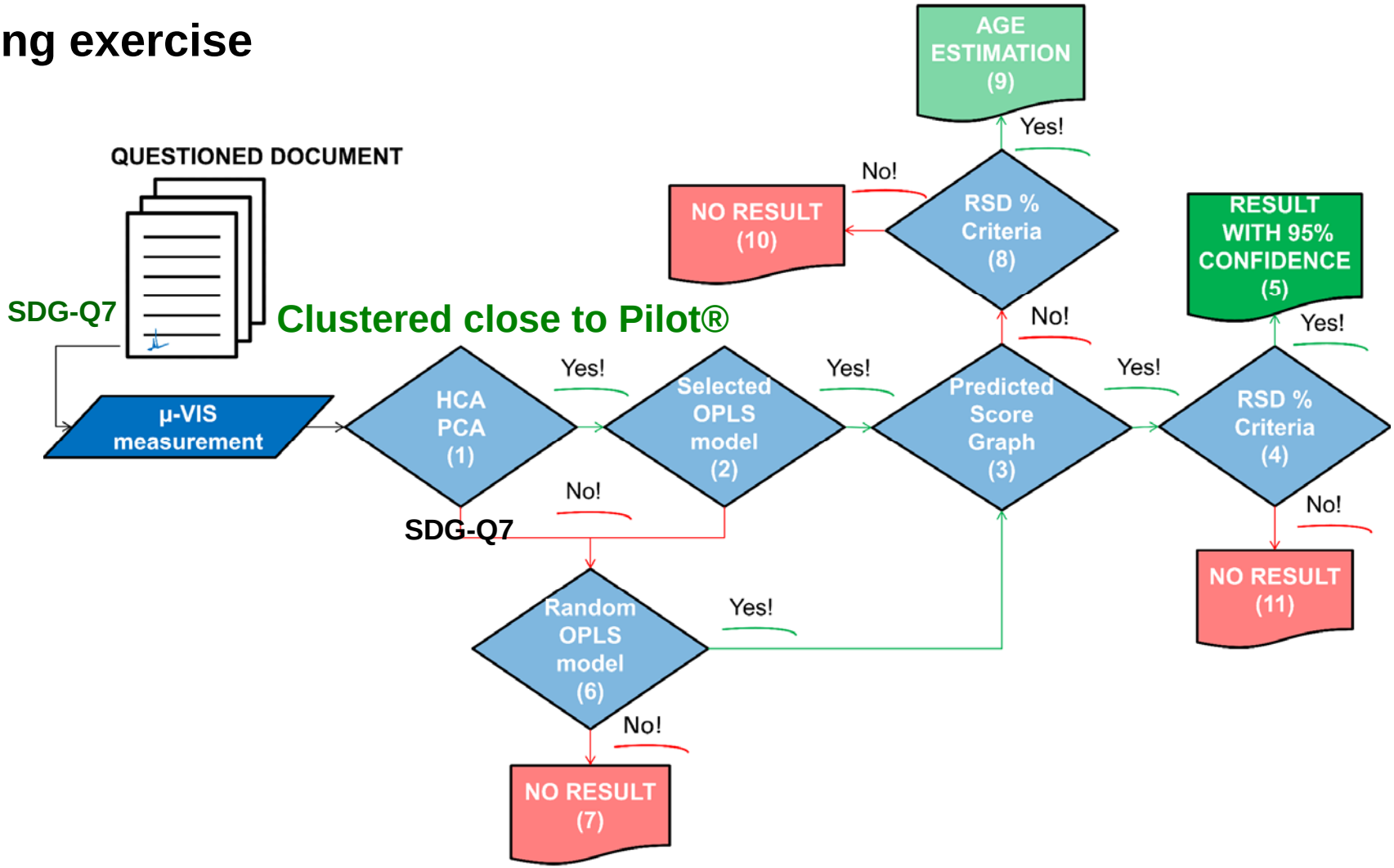




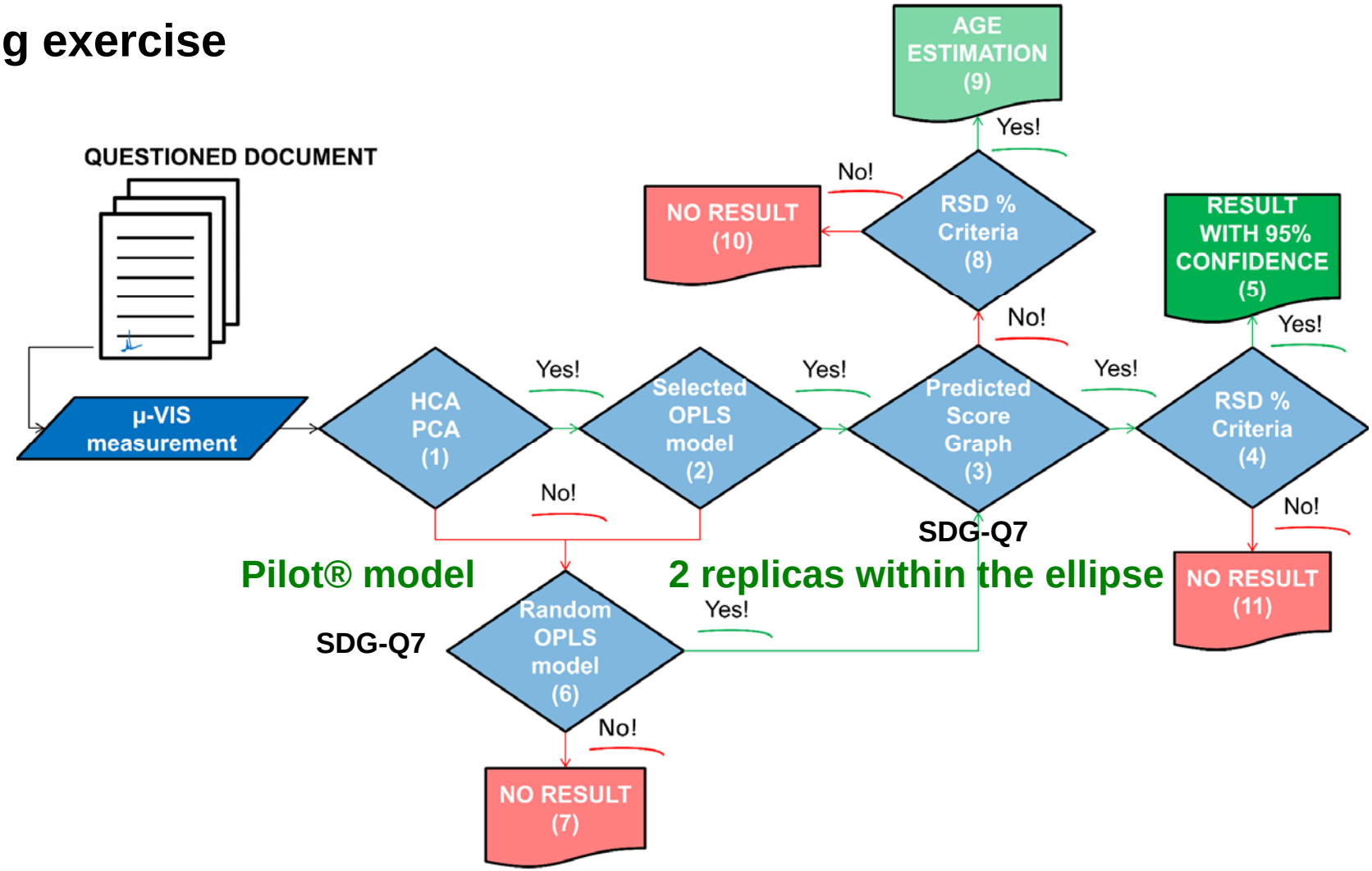
Blind testing exercise



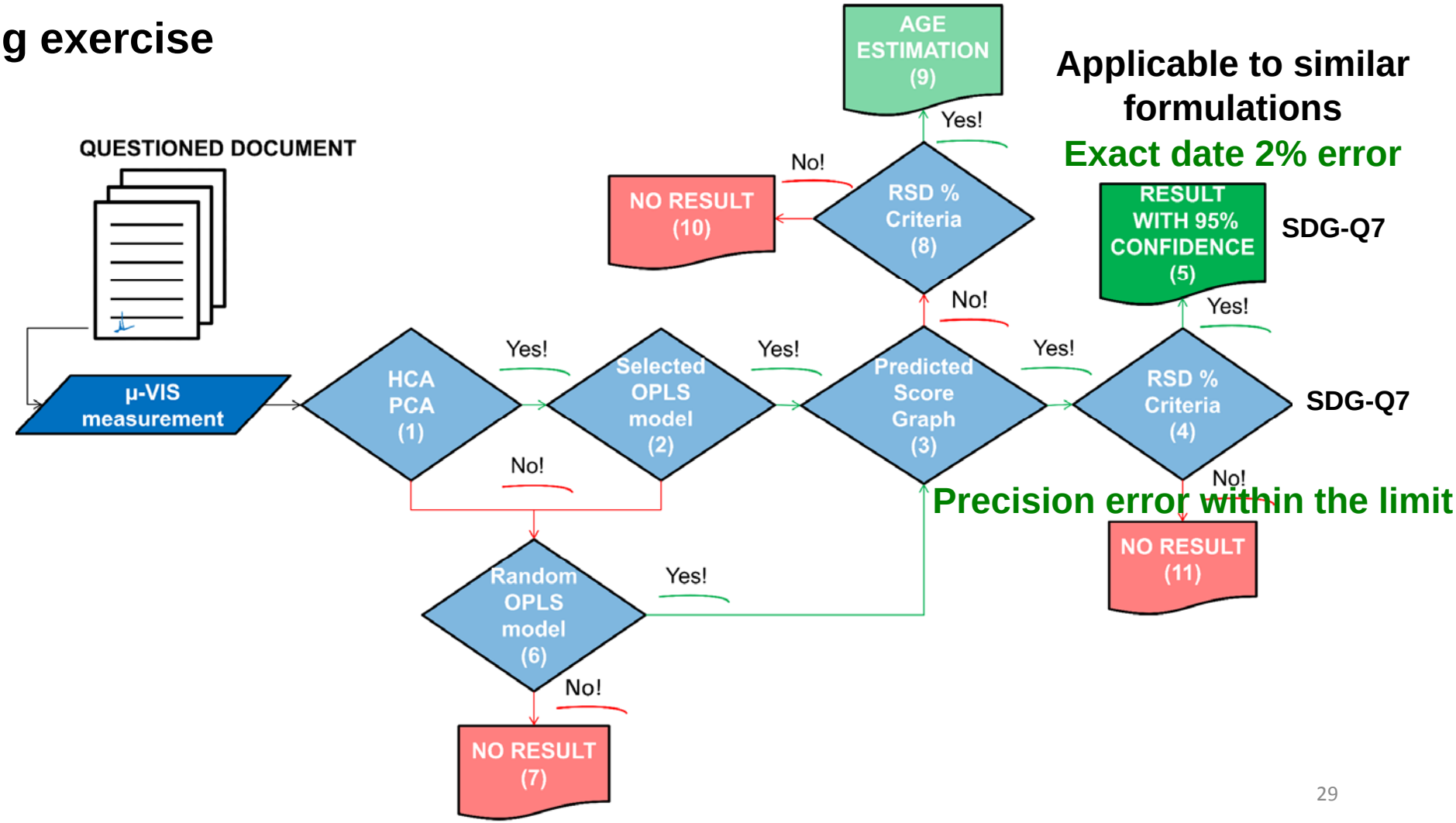
Blind testing exercise



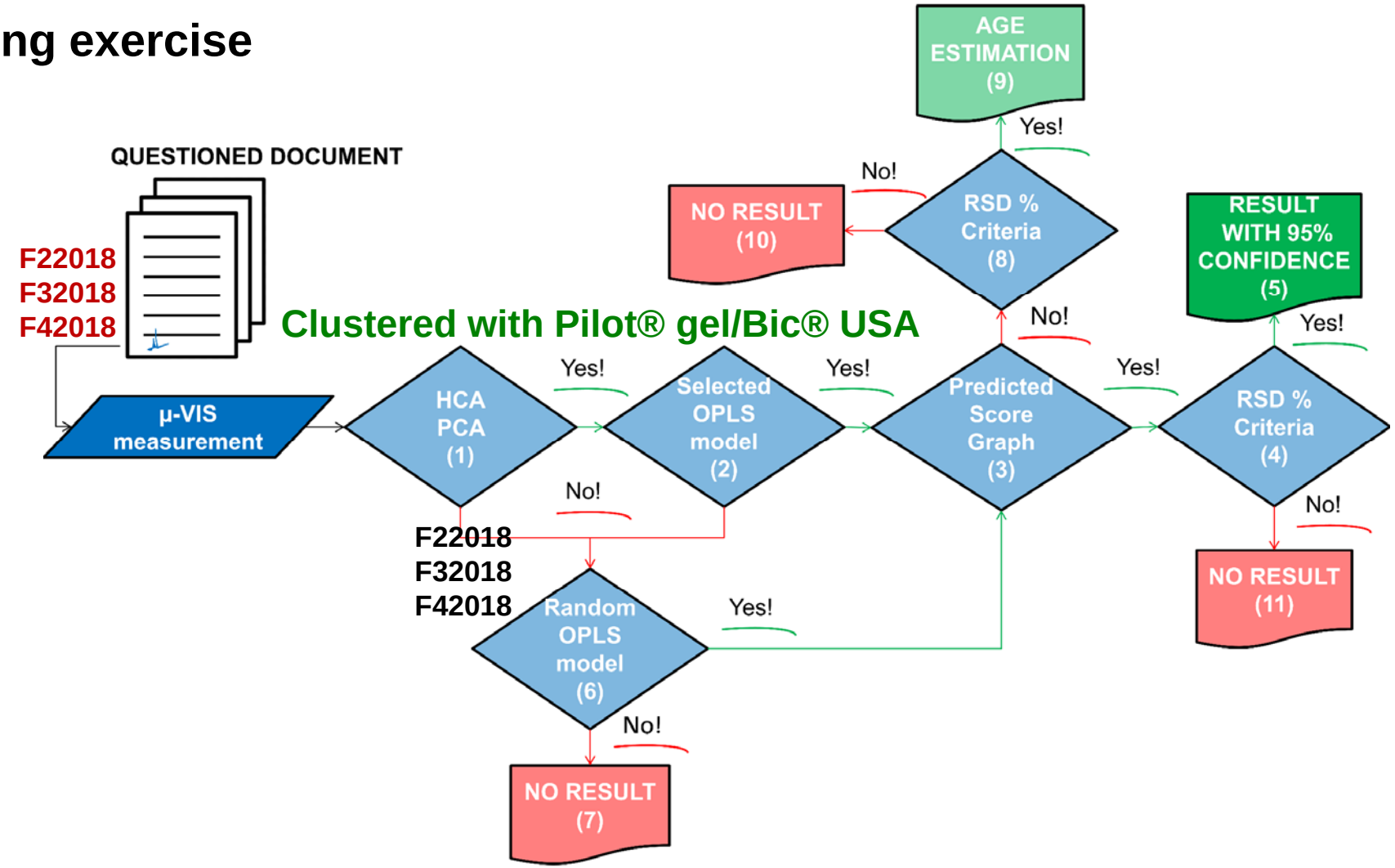
Blind testing exercise



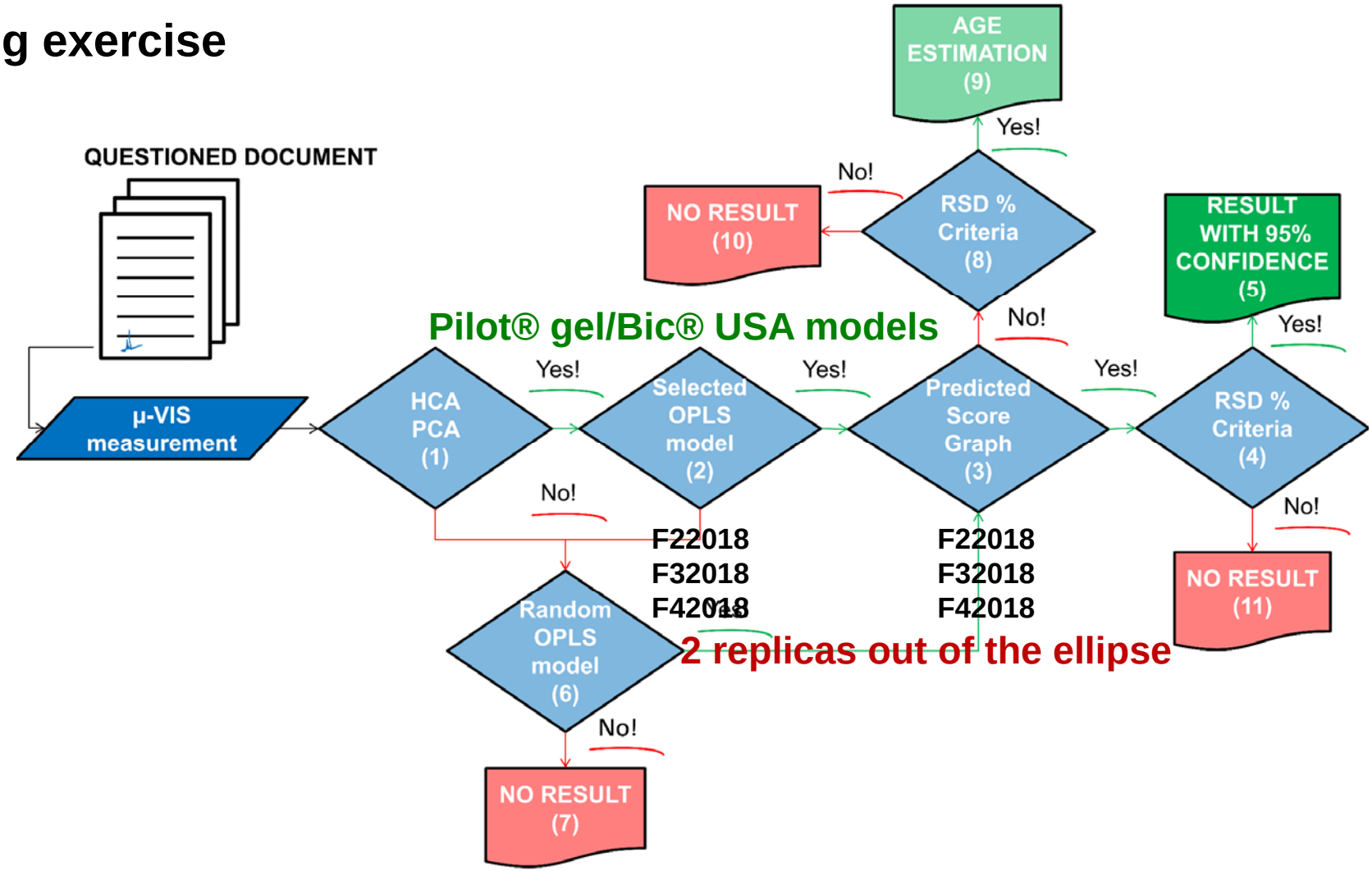
Blind testing exercise



Blind testing exercise



Blind testing exercise

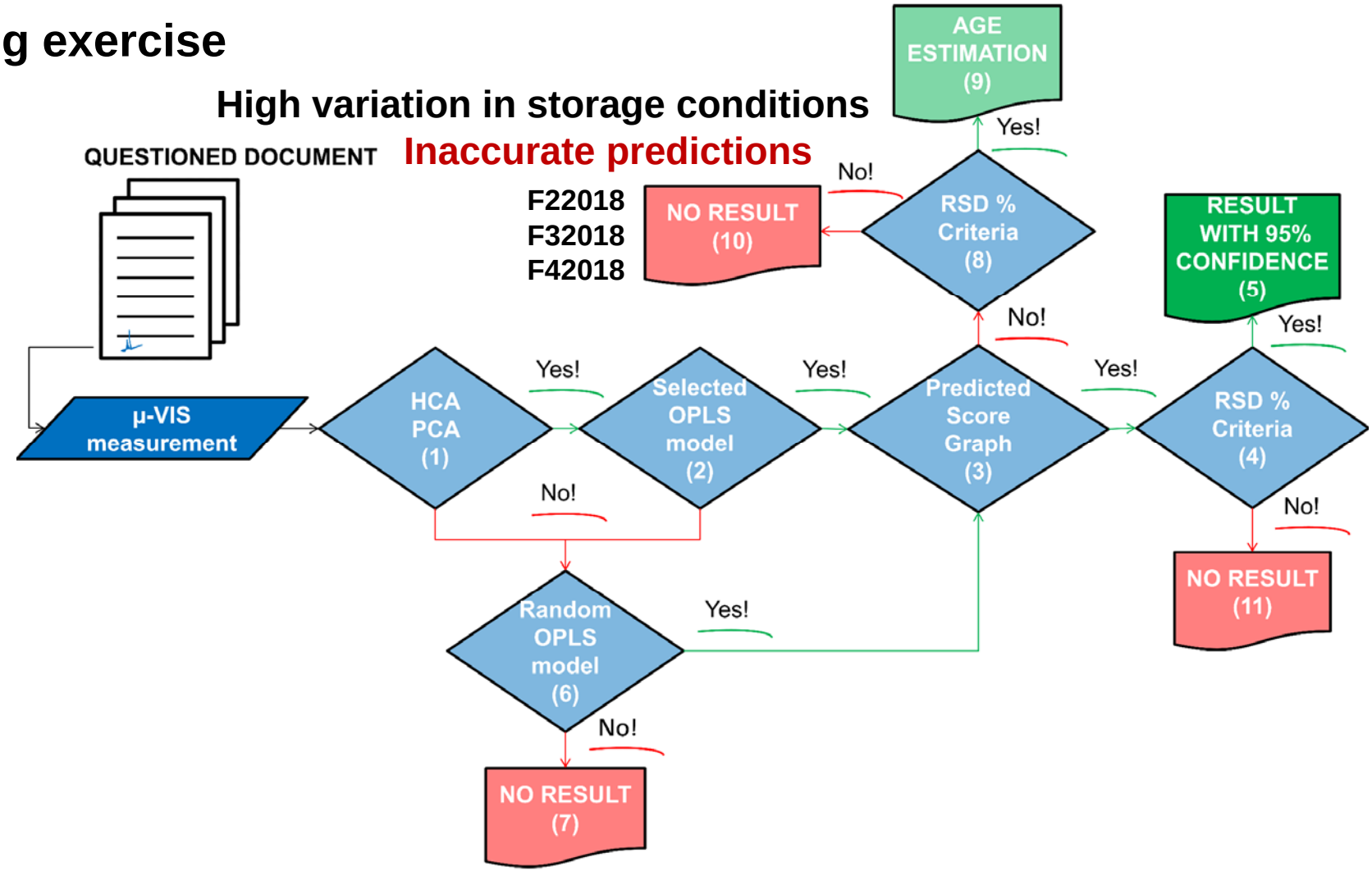


Blind testing exercise

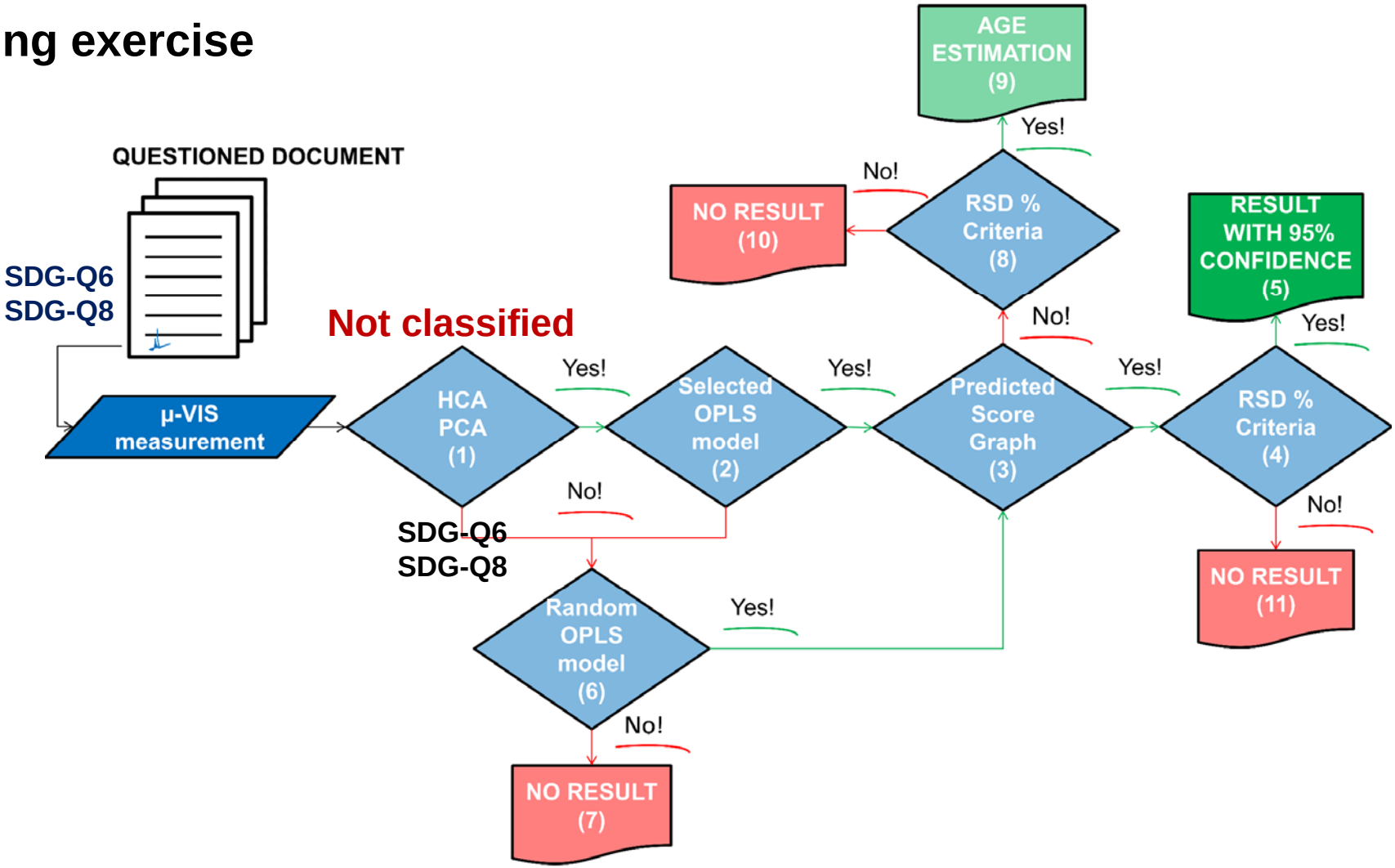
High variation in storage conditions

Inaccurate predictions

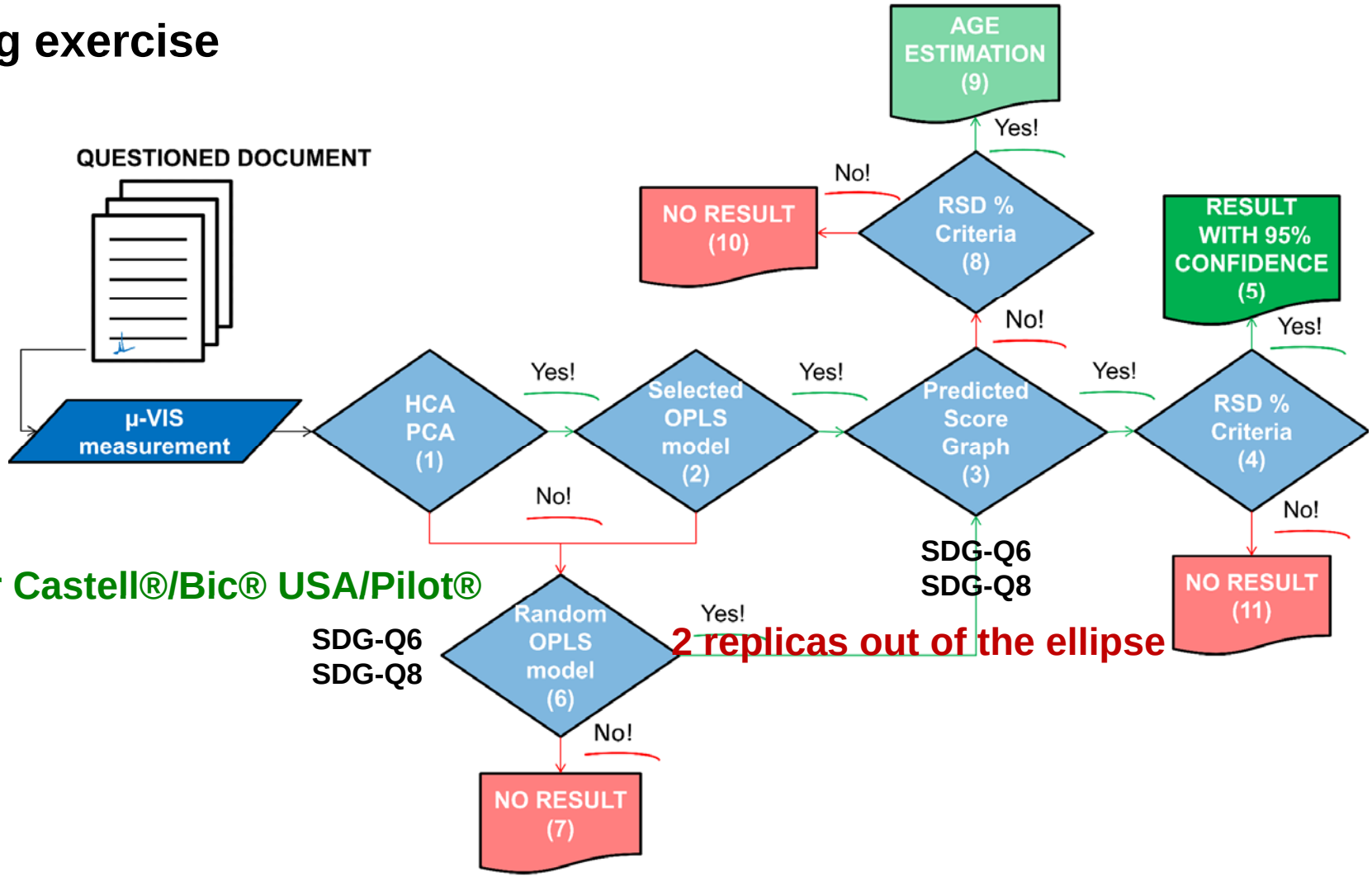
F22018
F32018
F42018



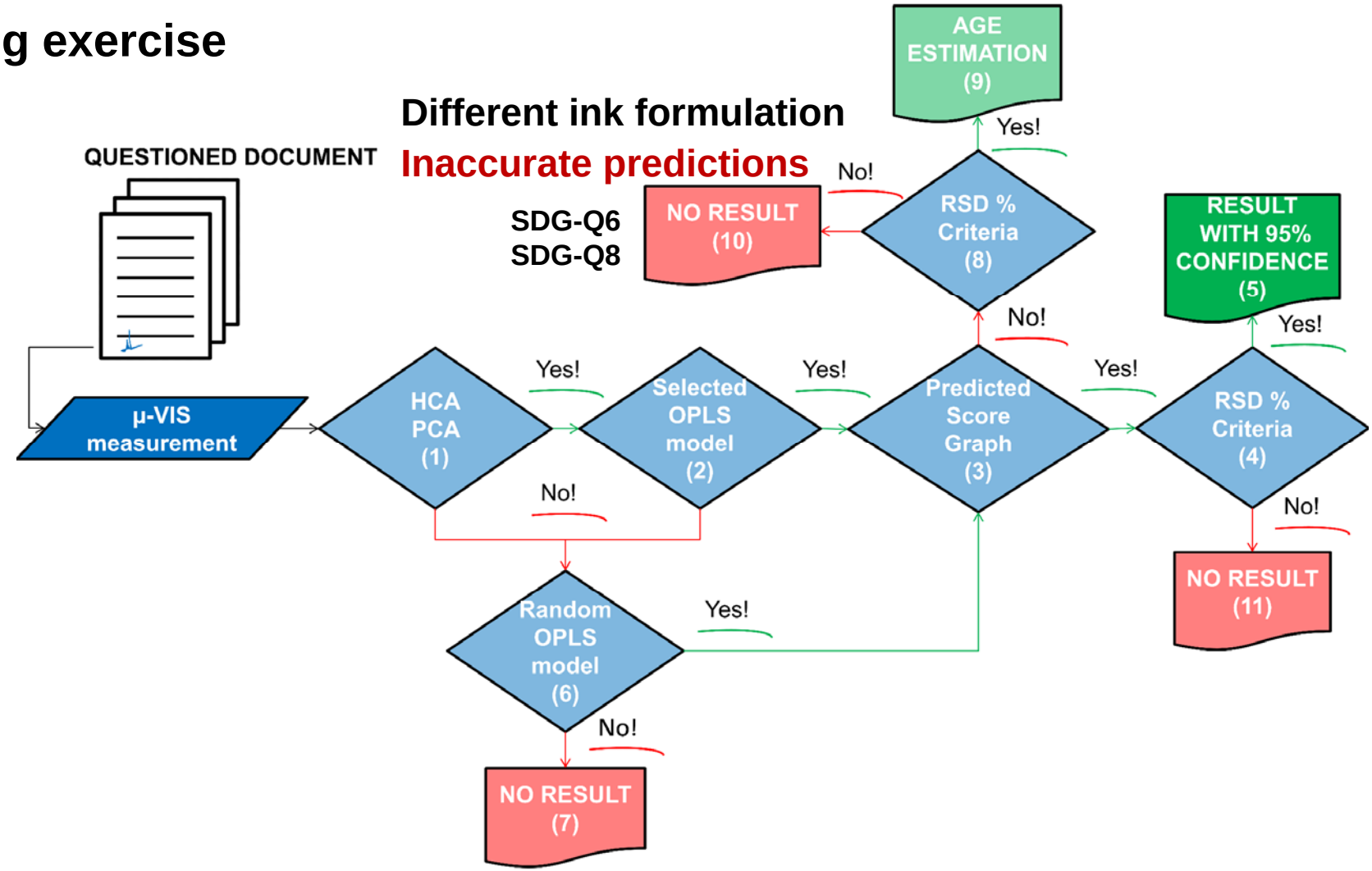
Blind testing exercise



Blind testing exercise



Blind testing exercise



Remarks

Exact date prediction with **95% confidence** whenever:

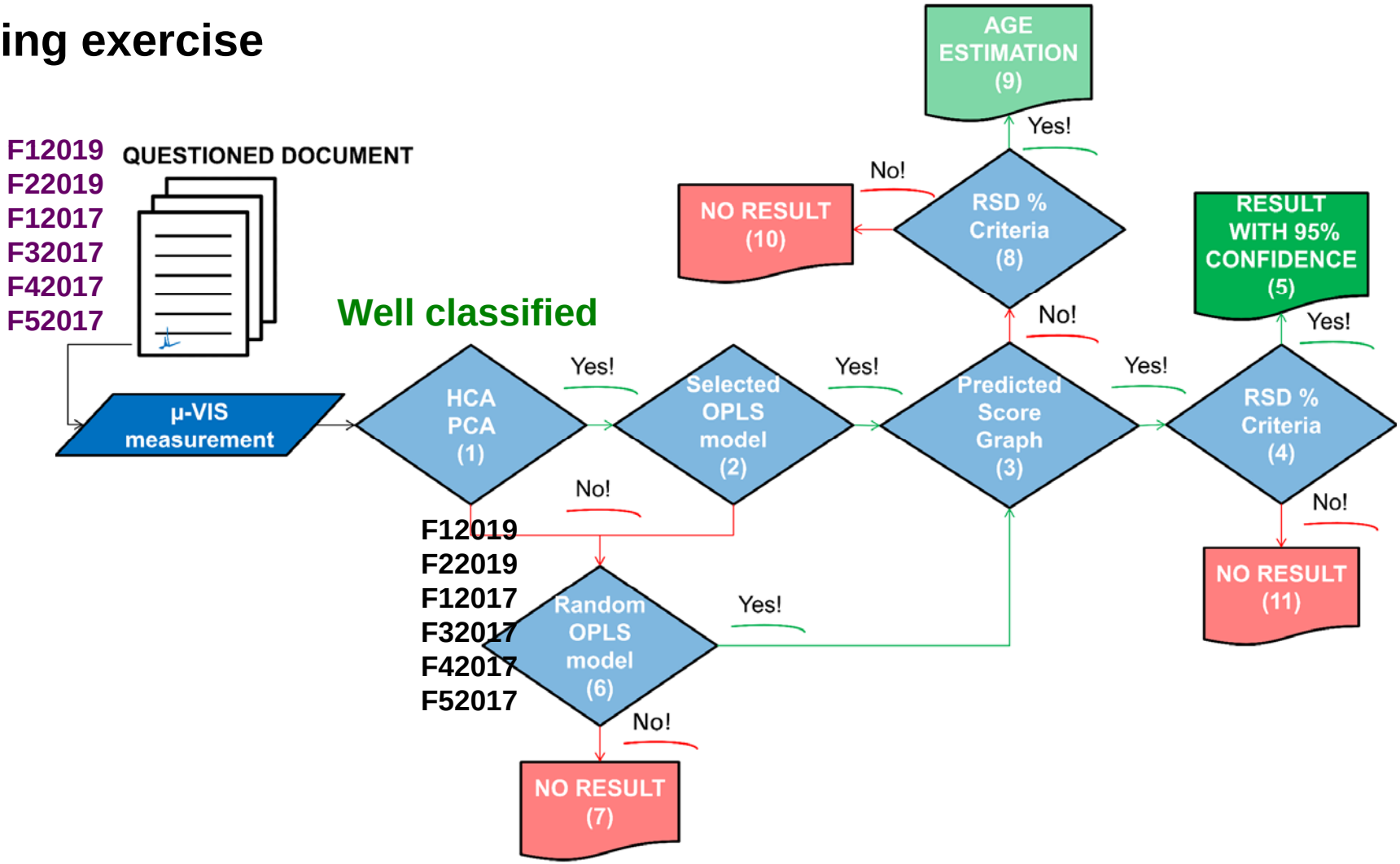
I. **Both replicas** lie **within the ellipse** of the predicted score plot.

II. The **age** of the ink is **within the time application range** of the corresponding OPLS model.

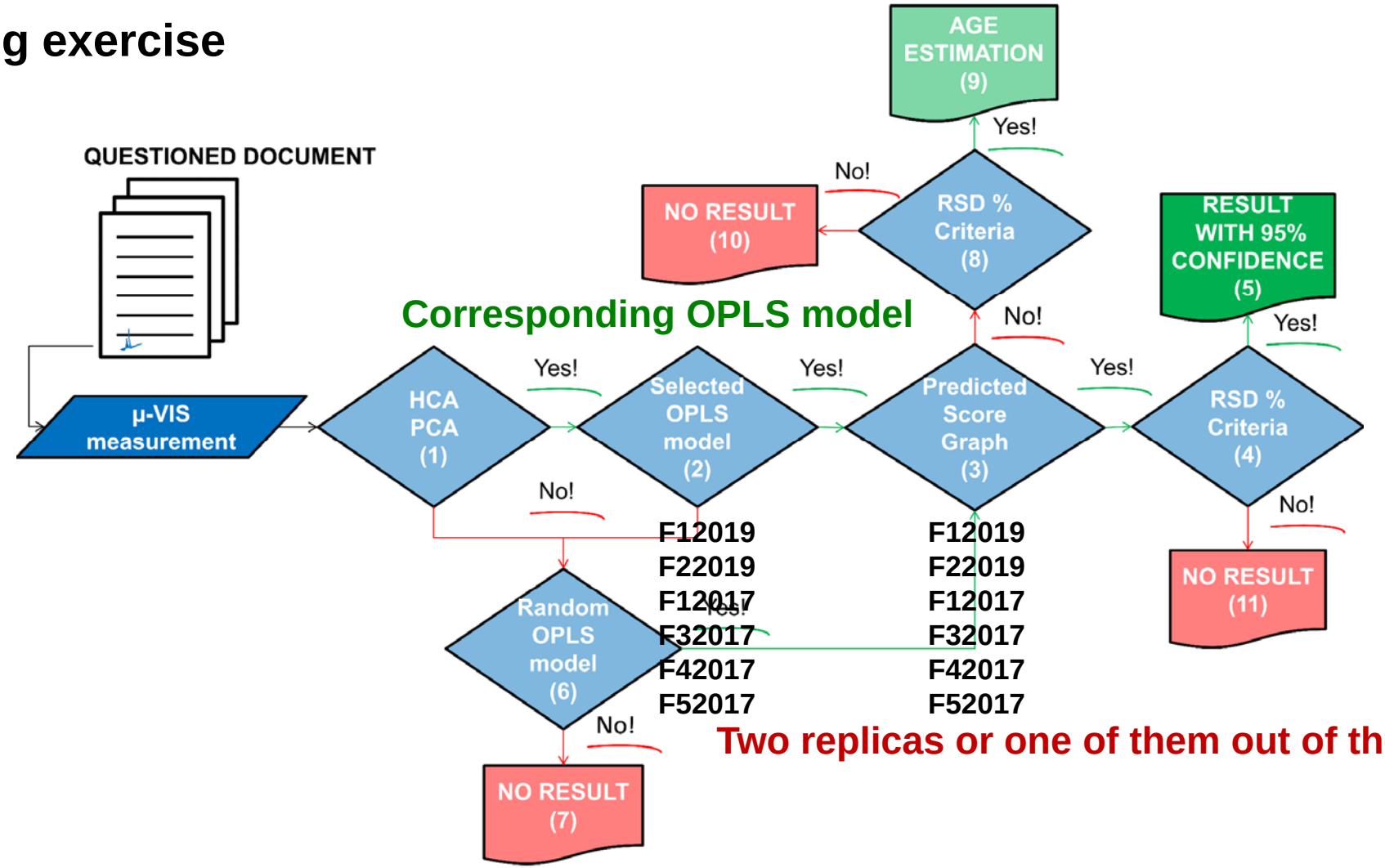
III. The ink is **well classified/clustered** into one of the classes/groups of pen brands studied.

$= / \approx$

Blind testing exercise



Blind testing exercise

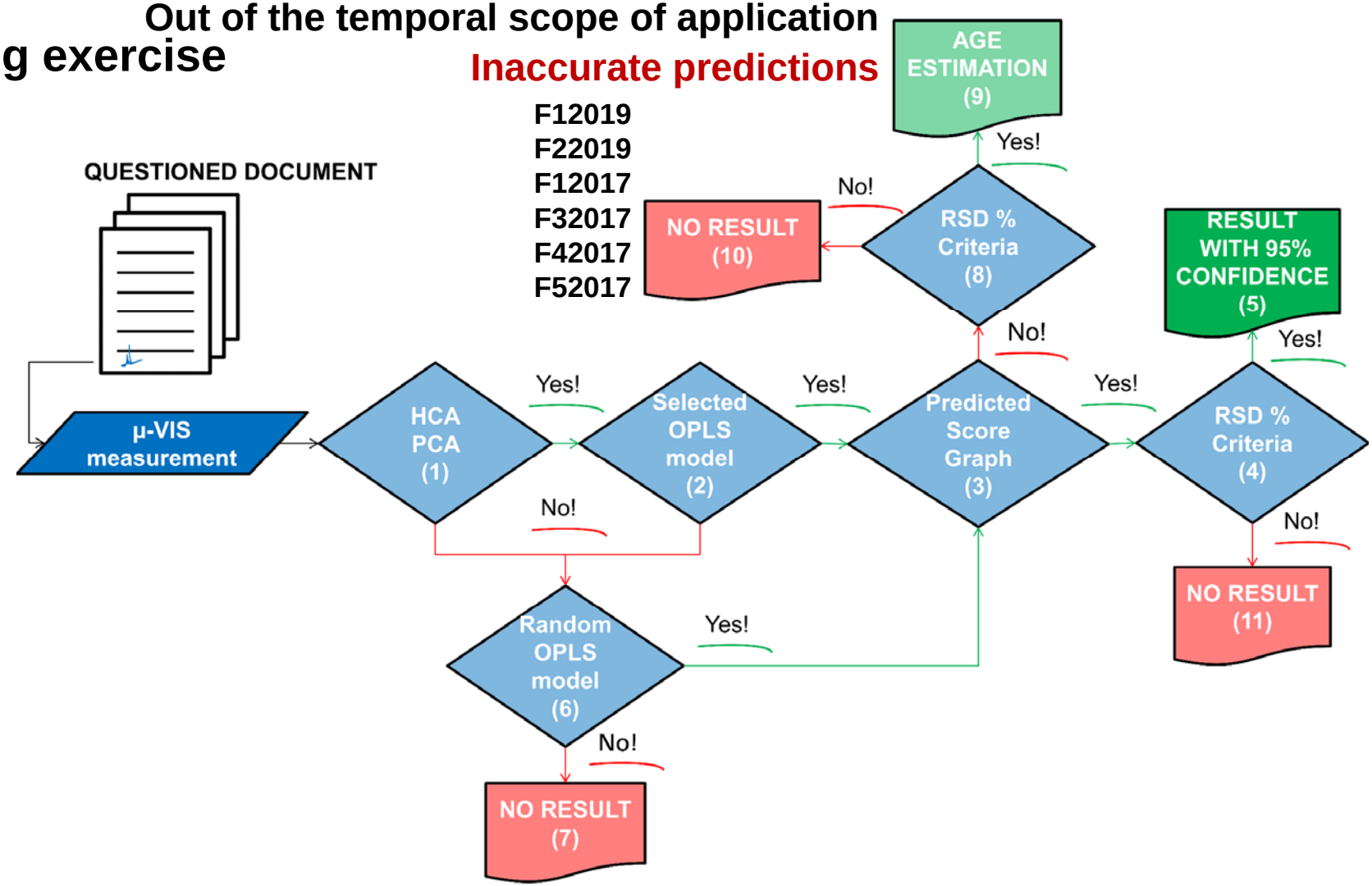


Blind testing exercise

Out of the temporal scope of application

Inaccurate predictions

F12019
F22019
F12017
F32017
F42017
F52017



Remarks

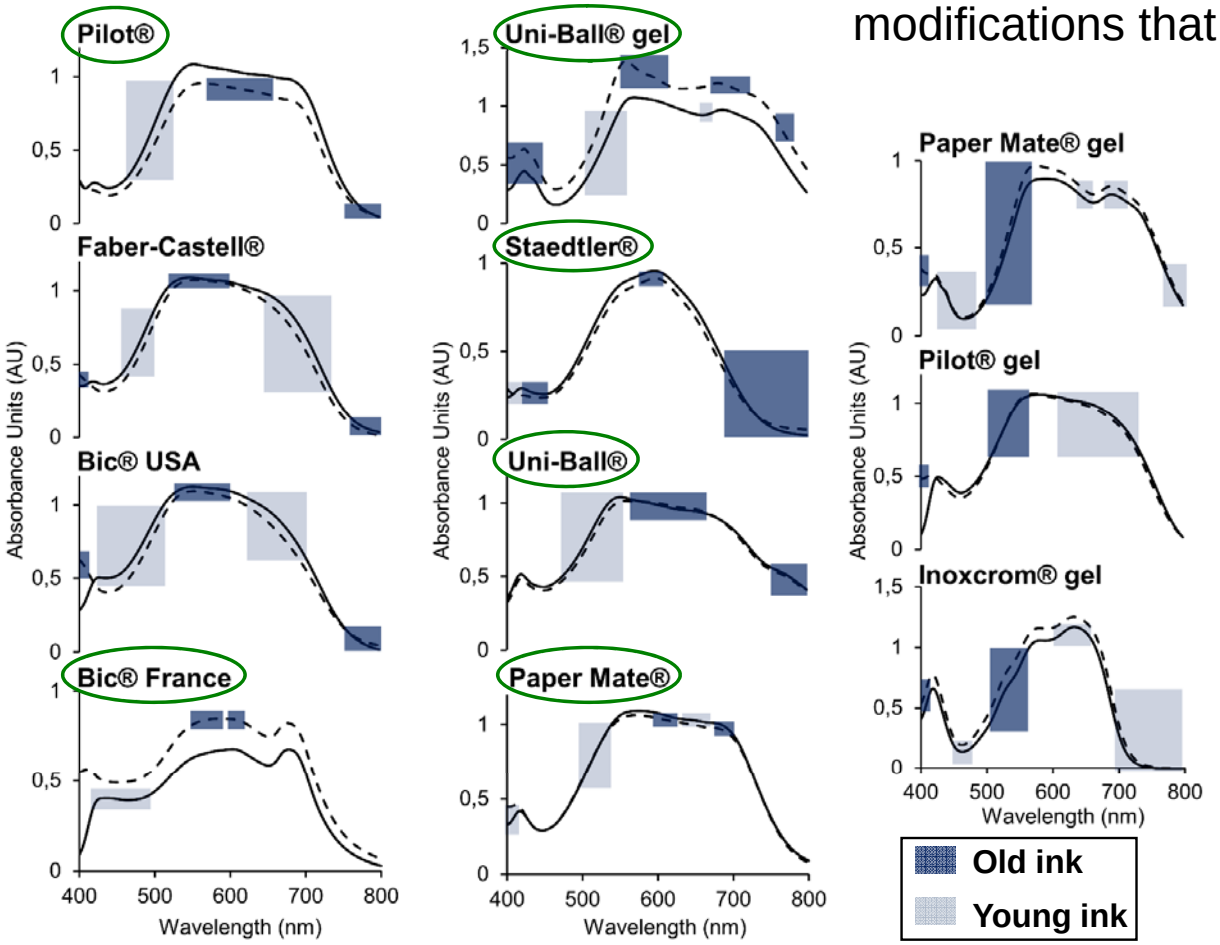
Ability to detect with **95% confidence** inks **out of their temporal application range** whenever:

I. The ink is **correctly clustered/classified** in the PCA/HCA model with one of the pen brands studied.

II. The two **ink replicas** or one of them falls **out of the ellipse** of the predicted score plot.

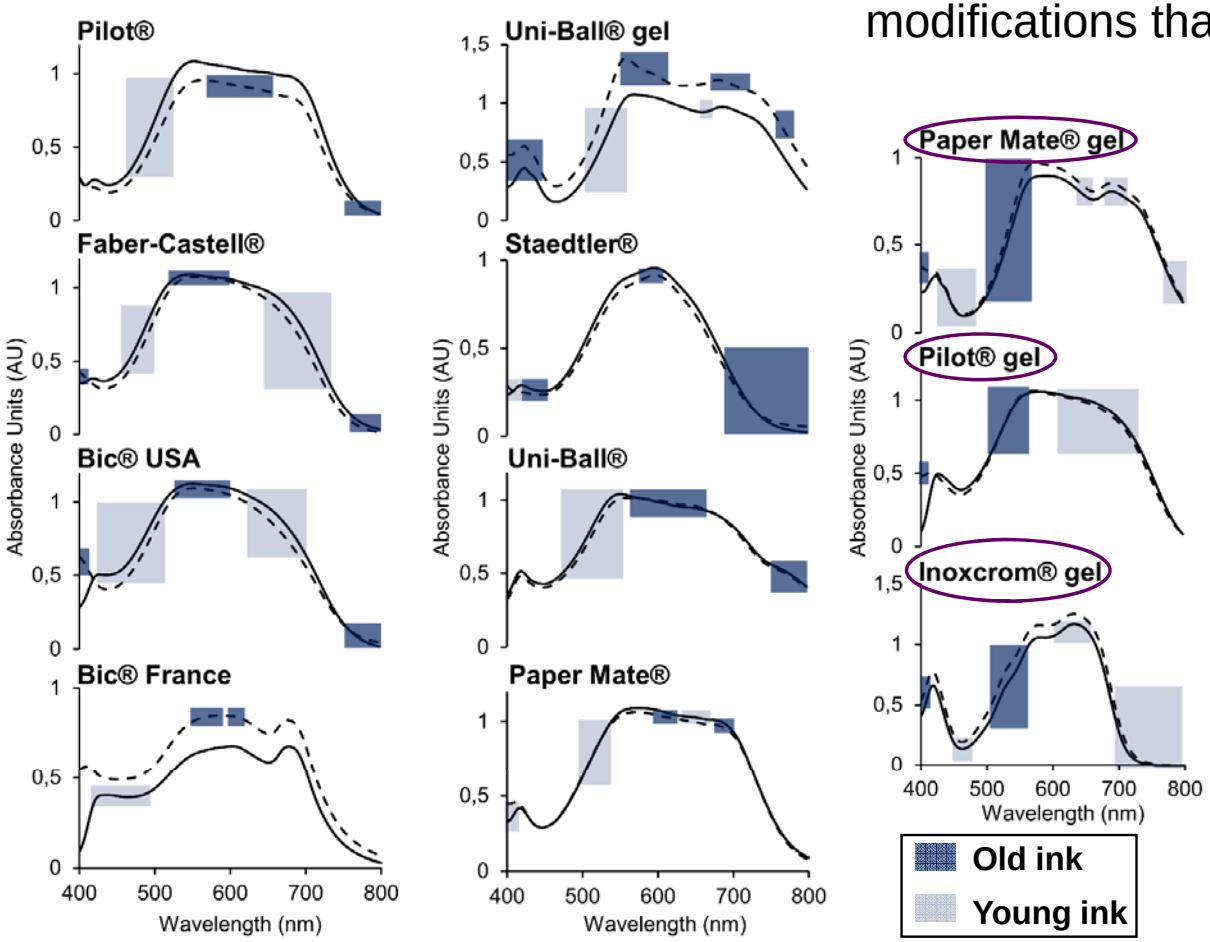
Loadings plot

Ink brands characterized over time by spectral modifications that **shift toward higher wavelengths**



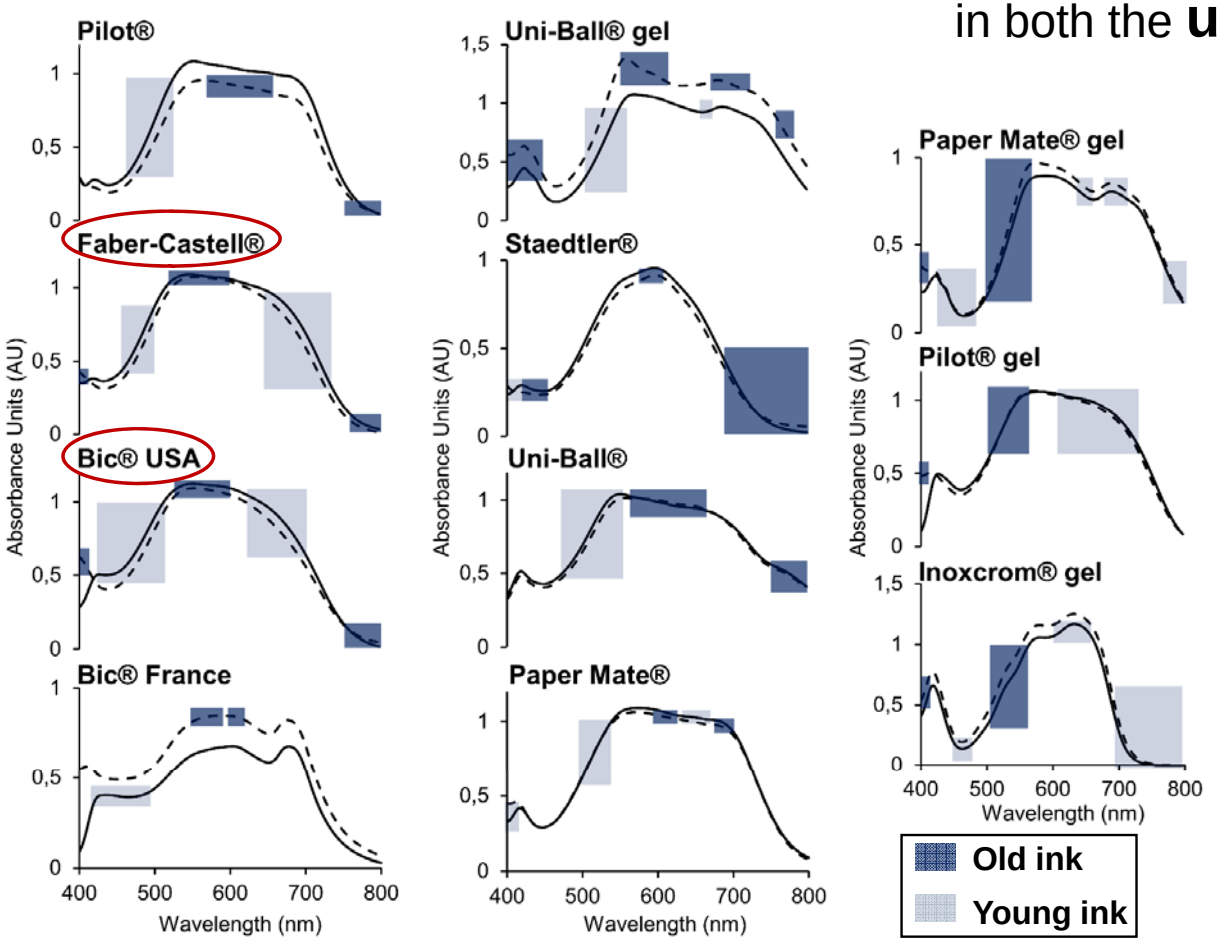
Loadings plot

Ink brands characterized over time by spectral modifications that **shift toward lower wavelengths**



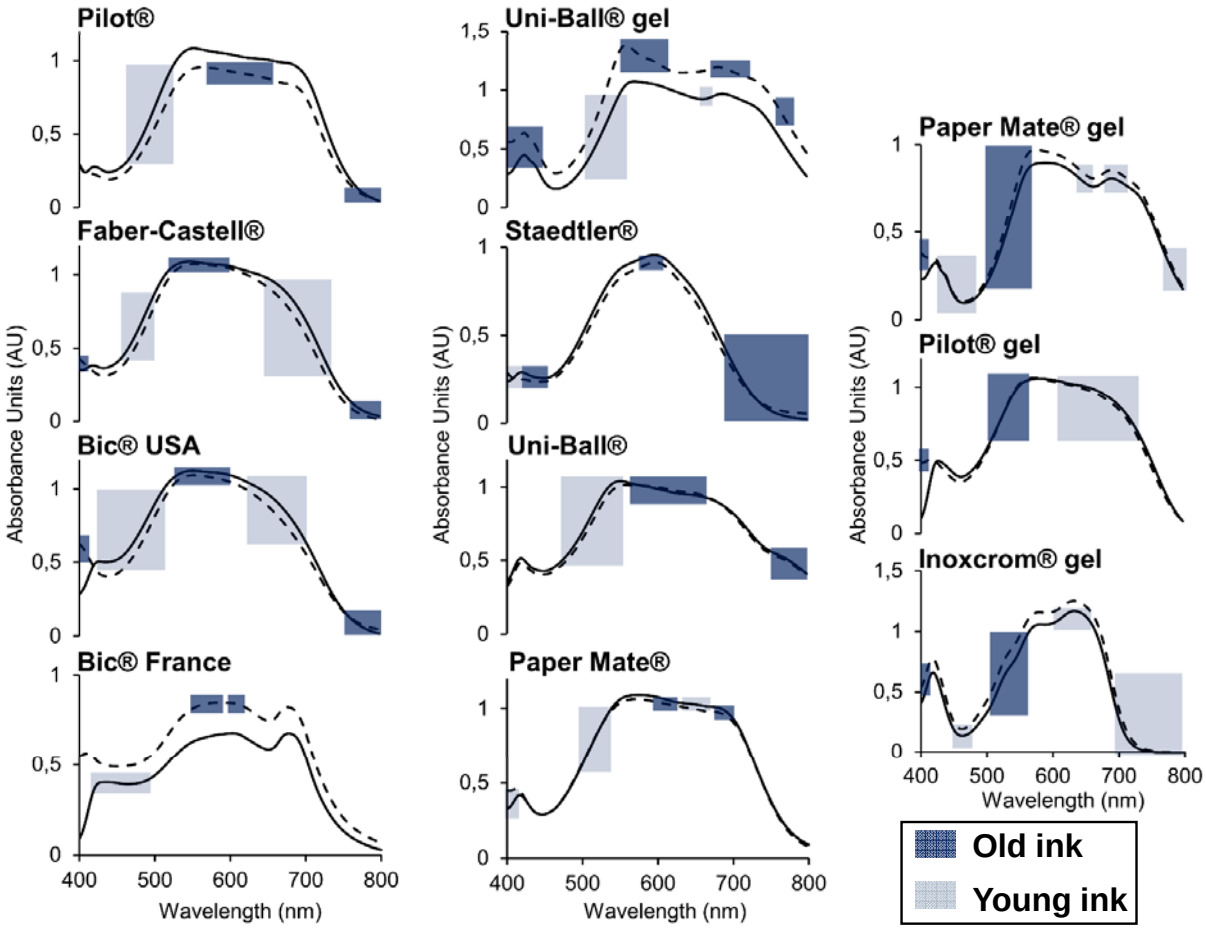
Loadings plot

Ink brands characterized over time by spectral modifications in both the **upper and lower wavelengths**



Loadings plot

Inability to obtain universal OPLS models



- **Multivariate regression methods** demonstrate great **versatile** of:
  - Application to **different forensic fields**.
  - Easily coupling with the most commonly used **analytical techniques**.

- The preparation of **custom-made** synthetic laboratory **samples** aged under **accelerated conditions**:
  - Overcome the **unavailability of samples**.
  - Enable **broad temporal scope** methodologies in a rapid manner.

- Multivariate regression **models** are easily **updated** as new samples are added.

- **New** chemometric-based dating **methodologies**:
  - Respond to **complex real-life claims** that conventional methodologies fail to cope with.
  - Improve and overcome the **disadvantages** or **objectives not reached** by their predecessors.

- The **statistical basis** of chemometrics provides **reliable** and **objective** age estimates.

- The **guideline** developed could be a valuable tool for use by the **forensic community**.

- **Testing** for the real **applicability** of the methodologies should involve the **participation** of experts.

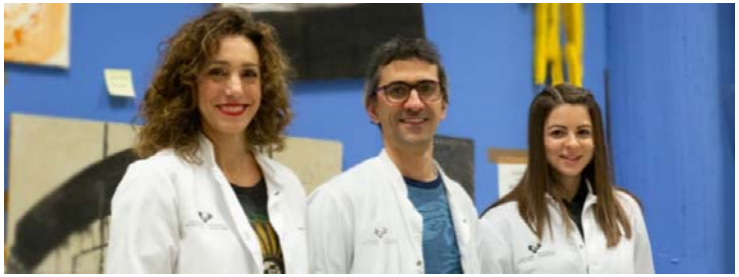
Team and collaborating academic and law enforcement institutions

Universidad del País Vasco · Euskal Herriko Unibertsitatea

FACULTY OF SCIENCE AND TECHNOLOGY
UNIVERSITY OF THE BASQUE COUNTRY

sgiker
Ikerkuntzarako Zerbitzu Orokorrak
Servicios Generales de Investigación

· U C · FCTUC FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE COIMBRA



Contact information:

luis.bartolome@ehu.eus laura.ortiz@ehu.eus