Version of Record: https://www.sciencedirect.com/science/article/pii/S1590865820304424
Manuscript_a55868f8c543591b86ac3e2e99cc608e

Identifying new potential genetic biomarkers for HELLP syndrome using massive parallel sequencing

Karen Jiménez [1†], Adrien Morel [1†], Laura Parada-Niño, María Alejandra González-Rodriguez [1], Stephanie Flórez [2], David Bolívar-Salazar [1], Silvia Becerra-Barona [3], Angel Aguirre-García [2], Tatiana Murcia [2], Luisa Fernanda Castillo [1], Carolina Carlosama, Javier Ardila [2], Daniel Vaiman [4], Norma Serrano [5] and Paul Laissue [1,4, 6*]

[1] Genetics Group. School of Medicine and Health Sciences. El Rosario University. Bogotá, Colombia.

[2] Departamento de Ginecología y Obstetricia. Hospital Universitario Mayor-Méderi. Bogotá, Colombia.

[3] Universidad Autónoma de Bucaramanga-UNAB, Colombia.

[4] Inserm U1016, CNRS UMR8104, Institut Cochin, équipe FGTB, 24, rue du faubourg Saint-Jacques, 75014 Paris, France.

[5] Research Centre, Fundación Cardiovascular de Colombia (FCV), Bucaramanga, Colombia.

[6] Orphan Diseases Group. Biopas Laboratoires. Bogotá, Colombia.

[†] **These authors contributed equally to this work.**

***Corresponding author:** Prof. Paul Laissue
Inserm U1016, CNRS UMR8104, Institut Cochin, équipe FGTB, 24, rue du faubourg Saint-Jacques, 75014 Paris, France.
E-mail: paullaissue@yahoo.com

© 2020 published by Elsevier. This manuscript is made available under the CC BY NC user license
https://creativecommons.org/licenses/by-nc/4.0/

Abstract

Background: Preeclampsia (PE) is a frequently occurring multisystemic disease affecting ~5% of pregnancies. PE patients may develop HELLP syndrome (haemolysis, elevated liver enzymes, and low platelet), a mother and foetus life-threatening condition. Research into HELLP's genetic origin has been relatively unsuccessful, mainly because normal placental function and blood pressure regulation involve the fine-regulation of hundreds of genes.

Objective: To identify new genes and mutations constituting potential biomarkers for HELLP syndrome.

Study design: The present case-control study involved whole-exome sequencing of 79 unrelated HELLP women. Candidate variants were screened in a control population constituted by 176 individuals. Stringent bioinformatics filters were used for selecting potentially etiological sequence variants in a subset of 487 genes. We used robust *in silico* mutation modelling for predicting the potential effect on protein structure.

Results: We identified numerous sequence variants in genes related to angiogenesis/coagulation/blood pressure regulation, cell differentiation/communication/adhesion, cell cycle and transcriptional gene regulation, extracellular matrix biology, lipid metabolism and immunological response. Five sequence variants generated premature stop codons in genes playing an essential role in placental physiology (*STOX1*, *PDGFD*, *IGF2*, *MMP1* and *DNAH11*). Six variants (ERAP1- p.Ile915Thr, ERAP2- p.Leu837Ser, COMT-p.His192Gln, CSAD-p.Pro418Ser, CDH1- p.Ala298Thr and CCR2-p.Met249Lys) led to destabilisation of protein structure as they had significant energy and residue interaction-related changes. We identified at

least two mutations in 57% of patients, arguing in favour of a polygenic origin for the HELLP syndrome.

Conclusion: Our results provide novel evidence regarding PE/HELLP's genetic origin, leading to new biomarkers, having potential clinical usefulness, being proposed.

Key words: Genetic biomarkers; HELLP syndrome; molecular medicine; protein modelling; whole-exome sequencing (WES).

Introduction

Preeclampsia (PE), clinically characterised by pregnancy-induced hypertension and proteinuria after the 20th week of gestation, is a frequently occurring multisystemic disease affecting ~5% of pregnancies [1,2]. Women affected by PE can suffer foetal intra-uterine growth restriction (IUGR) and have an increased risk of being affected by cardiovascular disease later in life [3,4]. It has been seen that ~15% of patients affected by severe PE may develop HELLP (haemolysis, elevated liver enzymes, and low platelet) syndrome, a mother and foetus life-threatening condition. Concerning PE *per se*, mechanisms have been described regarding PE, such as defective deep placentation, anticoagulant factor activation, immune response at the placental–maternal interface, increased endoplasmic and oxidative stress, having autoantibodies to type-1 angiotensin II receptor, intravascular inflammation, angiogenesis imbalance and endothelial dysfunction [5].

However, the exact mechanisms triggering PE developing into HELLP syndrome have not been completely elucidated. Several potentially PE-linked biochemical markers (e.g. PAPP-A, free CG-β, AFP, PP13, ADAM12, PlGF, VEGF, P-Selectin and inhibin A) have been studied for identifying those underlying possible usefulness in clinical environments [5]. For the moment, while some of these markers are indeed associated with preeclampsia, the effectiveness of those for predicting early PE is still disputable. Regarding genetics of PE, at least 15 PE loci have been mapped to date and candidate genes have been proposed, such as *ACVR2A*, *TNFSF13B*, *EPAS1* and *STOX1*[6,7]. The *STOX1* c.457T>C (p.Tyr153His) mutation has been functionally linked to PE aetiology, while *Stox1* overexpression in mice leads to PE, IUGR and cardiovascular disease [8–12]. The p.Ala356Gly deleterious mutation in Forkhead Box D1 (FOXD1), leading to *C3*

promoter transactivation disturbances, has been identified recently in a woman suffering PE and IUGR [13,14]. Some sequence variants in other genes (i.e. *VEGF*, *MMP9*, *SERPINEA8* and *MMP9*) seem to increase the risk of PE.

Regarding the HELLP syndrome, a few genetic studies have been undertaken and it has generally been accepted that, due to common molecular pathways, most PE-related genes may also contribute towards HELLP's aetiology.

However, research into PE/HELLP's genetic origin has been relatively unsuccessful, mainly because normal placental function and blood pressure regulation involve the fine-regulation of hundreds of genes. Furthermore, large pedigrees displaying PE/HELLP are rare, thereby generally hampering the mapping of genomic regions by classic genetic approaches. Whole-exome sequencing (WES) has been carried out in familial and isolated cases of PE which led to proposing some possible new genes and sequence variants related to the phenotype [15–18]. However, no large-scale genomic approaches have yet been adopted in patients with HELLP syndrome.

Here is presented the first WES study for patients affected by the HELLP syndrome, on a relatively large cohort (79 unrelated women) of Colombian origin. Stringent bioinformatics filters were used for selecting potentially etiological non-synonymous sequence variants in a subset of 487 genes. We used robust *in silico* mutation modelling for predicting the potential effect on protein structure. We identified 110 variants (87 different) in 67 genes related to several biological processes, such as angiogenesis/coagulation/blood pressure regulation, cell differentiation/communication/adhesion, cell cycle and transcriptional gene regulation, extracellular matrix biology, lipid metabolism and immunological response. Five sequence variants were particularly drastic as they generated premature stop codons in

genes playing an essential role in placental physiology (e.g. STOX1, PDGFD, IGF2, MMP1 and DNAH11). Six variants (ERAP1-p.Ile915Thr, ERAP2- p.Leu837Ser, COMT-p.His192Gln, CSAD-p.Pro418Ser, CDH1-p.Ala298Thr and CCR2-p.Met249Lys) led to potential destabilisation of protein structure as they had significant energy and residue interaction-related changes. We identified at least two mutations in 57% of patients, arguing in favour of a polygenic origin for the HELLP syndrome.

Taken together, our results provide novel evidence regarding PE/HELLP's genetic origin, leading to new biomarkers, having potential clinical usefulness, being proposed.

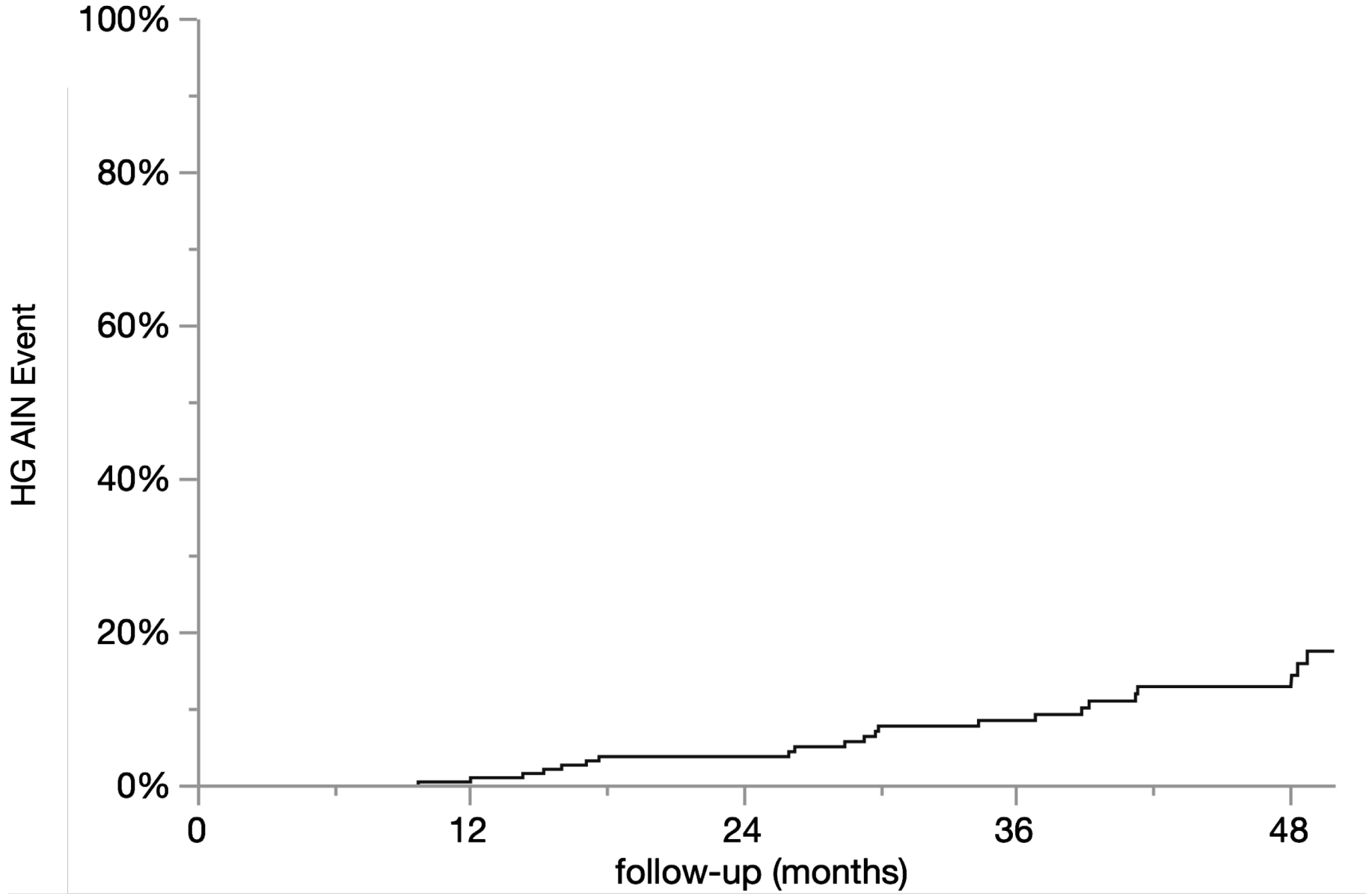
Materials and Methods

Women affected by HELLP syndrome and control population

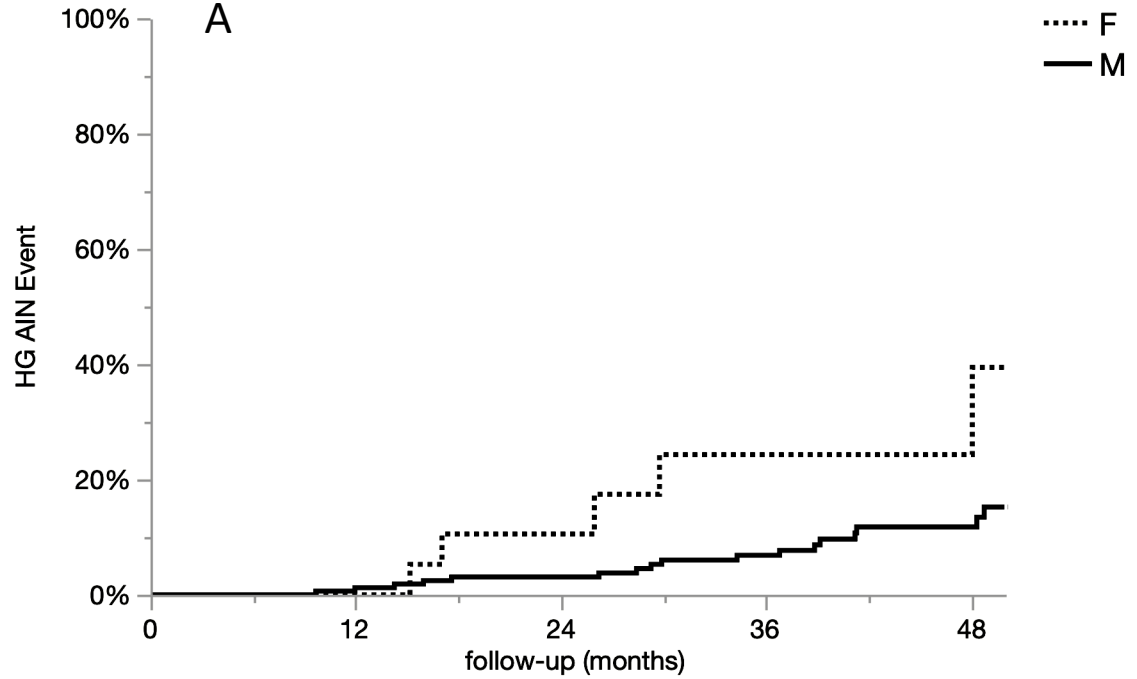
The study involved selecting 79 unrelated women affected by HELLP syndrome for WES sequencing. The patients (all from Colombia) are part of the GenPE study (Genetics and Pre-eclampsia), large, multi-centre, case-control study conducted between December 2000 and February 2012 in eight Colombian cities (www.genpe.org) or the Hospital Universitario Mayor-Méderi (Bogotá, Colombia). The **Supplementary Table 1** includes the clinical characteristics of the HELLP patients.

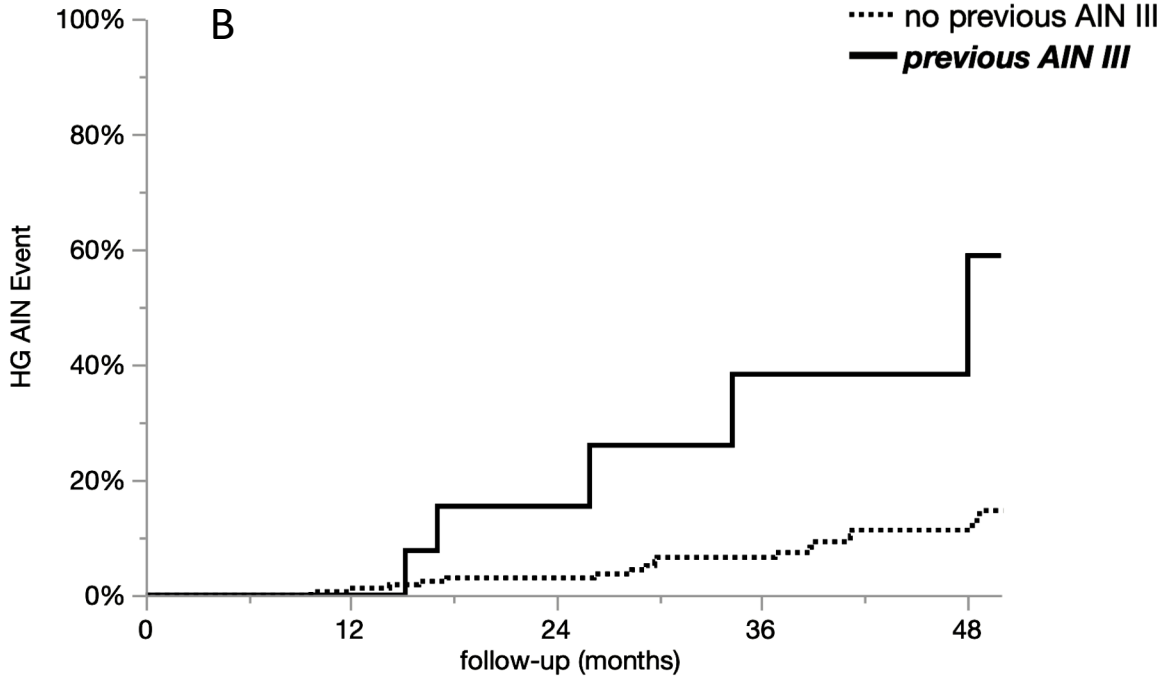
To calculate the sample size, we have first explored a Colombian health database (named SIVIGILA) that records numerous epidemiological events (http://www.ins.gov.co/buscador). We found that 22,915 women were affected by extreme gestational morbidity, 44.2% of which were suffered severe preeclampsia (PE). Thereafter, based on the reported frequency of HELLP (~20%) we estimated that 2028 should be affected by the disease. This number was taken as the total population. The

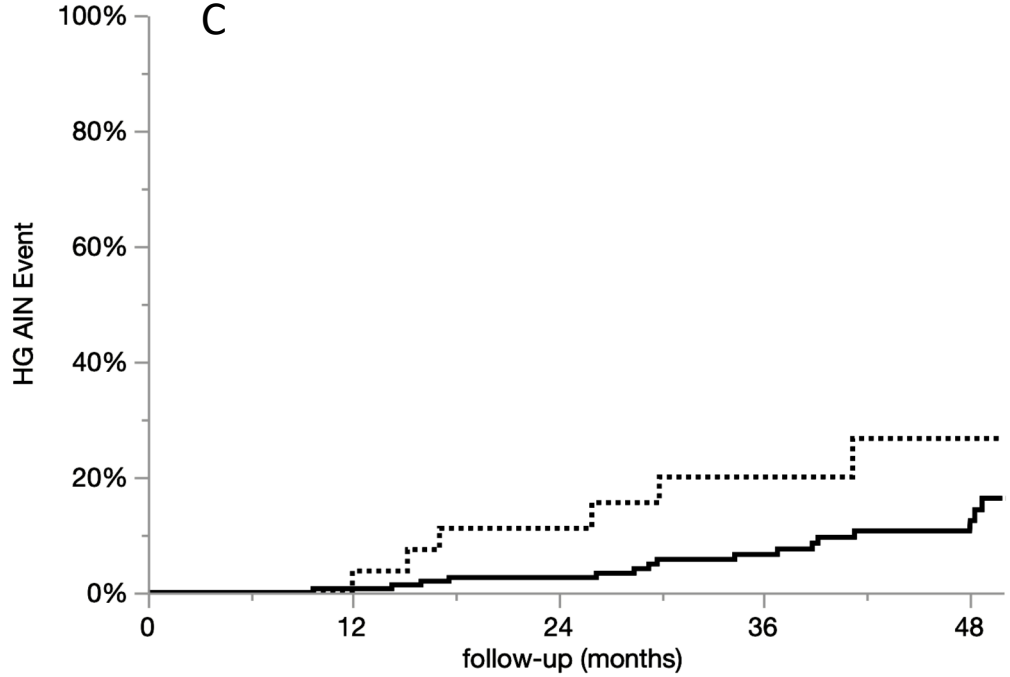
159 final calculation (5% proportion), which was performed considering a precision of 5%
160 and a confidence interval of 95%, led us to establish a sample size of 70 patients.
161 The Tennessee Classification System diagnostic criteria for HELLP syndrome [19] were
162 used for selecting patients. The main diagnostic laboratory criteria were thus: ≤
163 100,000/L platelets, ≥ 70 IU/ AST L and ≥ 600 IU/L LDH. This classification system
164 rates HELLP syndrome as being complete (three criteria) or incomplete (one or two
165 criteria). Candidate variants were screened using WES data from 176 Colombian women
166 (control group, C) over 50 years old having at least one child and lacking antecedents of
167 medical complications during pregnancy. Individuals having antecedents of hypertensive
168 disorders were excluded as controls. The Universidad del Rosario (Code: DVN-021-1-
169 063, May 2017) Ethics' Committee approved the clinical study and all experimental
170 procedures. Clinical investigation followed Declaration of Helsinki guidelines (1975, as
171 revised in 1996). All individuals had signed informed consent forms.
172
173 ***WES, bioinformatics filtering of variants, Sanger sequencing and interactome analysis***
174 Genomic DNA was extracted from all patients' blood samples, using standard
175 procedures. Details regarding NGS experiments and primary bioinformatics analysis have
176 been included as **Supplementary methods**. The detected sequence variants were filtered
177 in parallel using R (v3.5.1; www.r-project.org), Excel (Microsoft Office 365) and VarSeq
178 (Golden Helix Inc., Bozeman, USA) software. A subset of 487 candidate genes (named
179 HELLP-487) was created by screening several databases and used to filter rare (MAF
180 <0.005) variants evoking potentially harmful effects (**Supplementary methods,**
181 **Supplementary Table 2**). **Figure 1** illustrates the bioinformatics pipeline used for
182 candidate variants´ filtering. We screened these variants in the C group of women. All
183 selected variants were verified (n=118) by Sanger sequencing. The primer sequences and

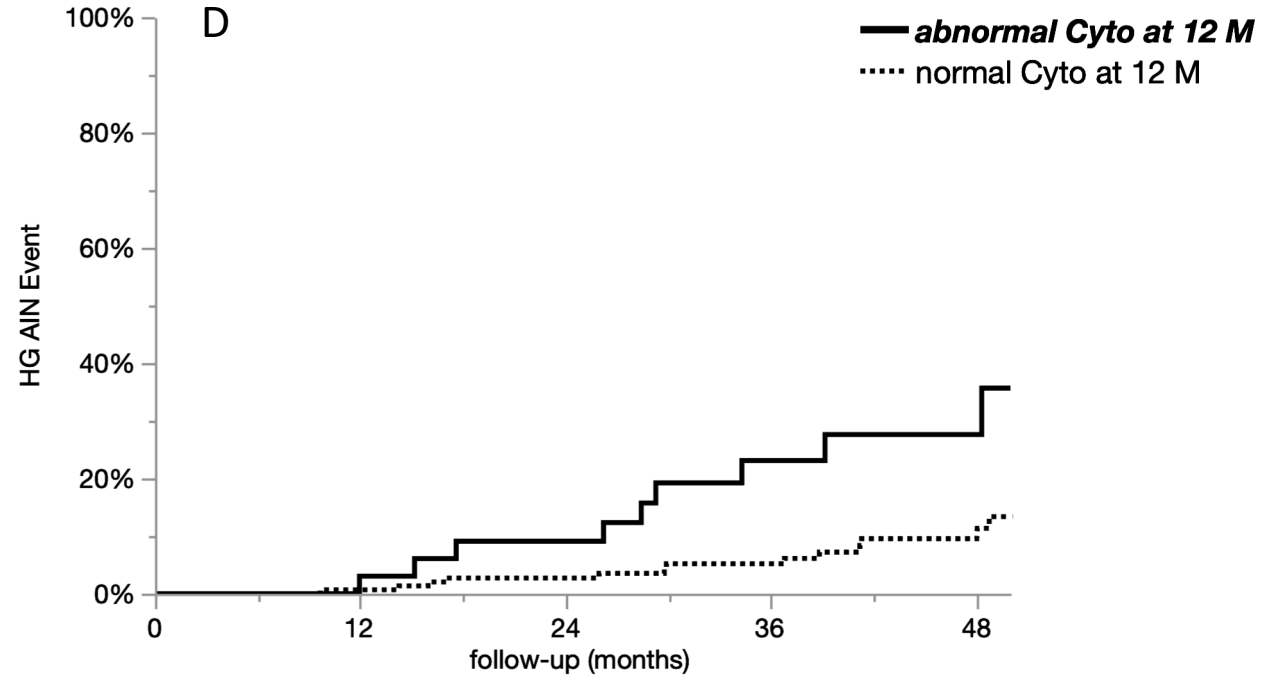


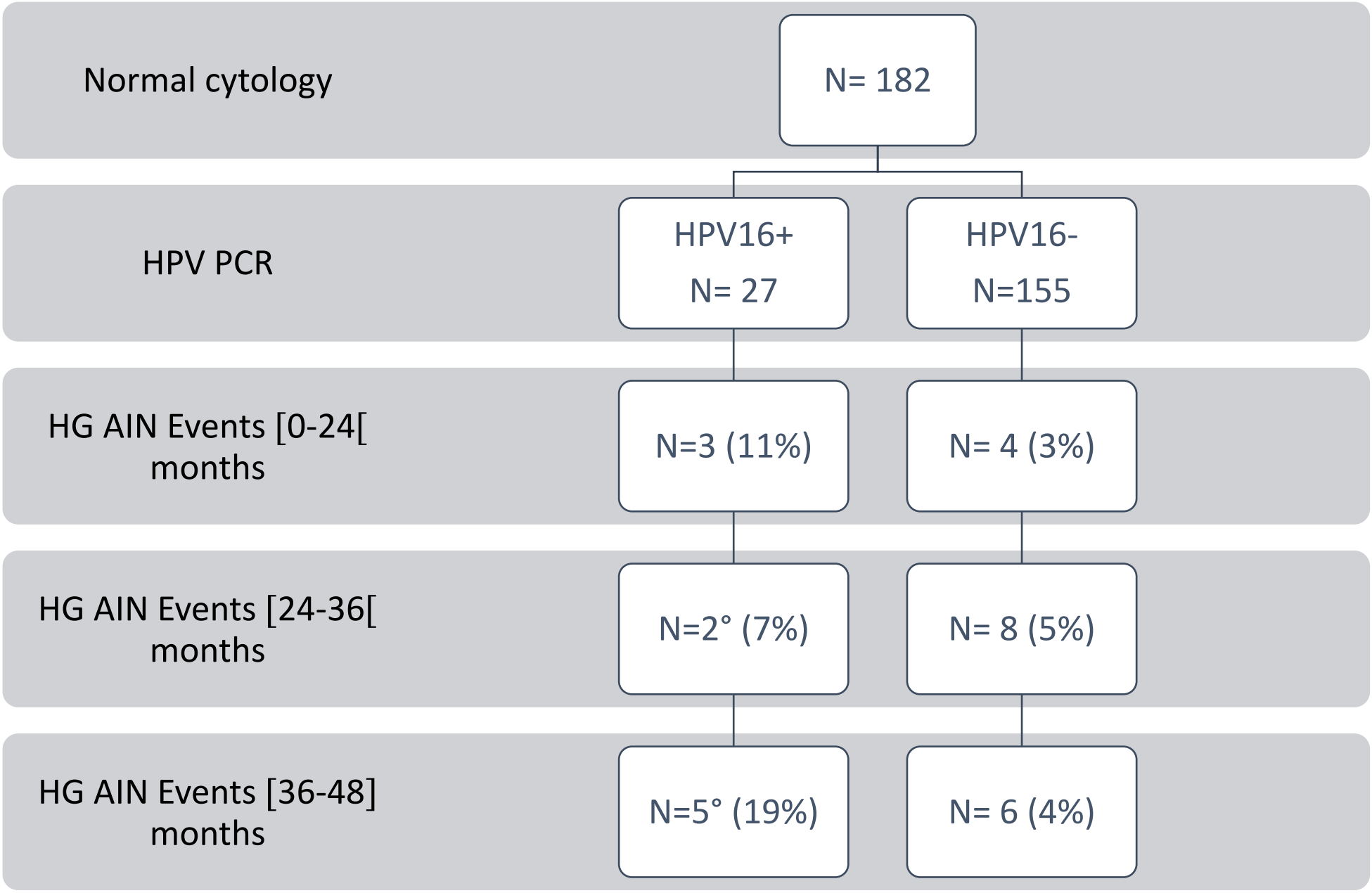
N at risk:	182	182	166	120	79


A
F
M
100%
80%
60%
40%
20%
0%
HG AIN Event
0
12
24
36
48
follow-up (months)


B
no previous AIN III
previous AIN III
100%
80%
60%
40%
20%
0%
HG AIN Event
0
12
24
36
48
follow-up (months)


C
No HPV16
HPV 16
100%
80%
60%
40%
20%
0%
HG AIN Event
0
12
24
36
48
follow-up (months)


D
abnormal Cyto at 12 M
normal Cyto at 12 M
100%
80%
60%
40%
20%
0%
HG AIN Event
0
12
24
36
48
follow-up (months)



° one invasive cancer

Table 1: Patient characteristics and anal events

Item	N	N (%) or mean (SD)
Age (years)	182	48.1 (11)
Sex (F/M)	182	19/163 (10/90)
Past history and comorbidities		
HPV-related conditions	182	77 (42)
Treatment of high-grade anal intraepithelial lesions	182	13 (7)
Cervical high-grade lesion or invasive cancer	19	2 (11)
Sexually transmitted diseases (other than HPV)	182	47(28)
HIV infection	182	147 (81)
HAART	147	124 (84)
Nadir CD4 count (cell/mm^3)	82	302 (214)
MSM	163	126 (77)
Tobacco consumption (yes/ex/no)	181	29/36/116 (16/20/64)
Tobacco exposure (months)	60	331(128)
Body mass index (kg/m^2)	177	23.5 (4)
BMI>30	177	12 (7)
Diabetes mellitus	182	1 (1)
Immunocompromised host (HIV exclusion)	180	2 (1)
Hysterectomy for cancer	19	1 (5)
Crohn's disease	182	1 (1)
HPV genotypes at baseline		
Number of HPV genotypes per patient	182	1.4 (1.6)
HPV16	182	27 (15)
High-risk HPV	182	81 (45)
Events within the study period		
Follow-up duration (months)	182	41.2 (10.7)
High-grade AIN/invasive cancer during the follow-up	182	28/2 (15/1)
AIN treatments during follow-up	182	19 (10)

Table 2: Events within the study period

Follow-up (months)	[0-12[	[12-24[	[24-36[	[36-48]
Pathology	182	166	120	79
Abnormal cytology	30 (17%)	54 (33%)	34 (28%)	35 (44%)
New HSIL	3 (2%)	4 (2%)	3 (3%)	8 (10%)
New HG AIN	2 (1%)	5 (3%)	10 (8%)	11 (14%)
Virology	182	161	114	73
New HPV16	27 (15%)	6 (4%)	5 (4%)	8 (11%)
HPV16+	34 (19%)	33 (21%)	28 (25%)	26 (36%)
HR HPV+	87 (48%)	80 (48%)	66 (58%)	48 (66%)

Table 3: Factors associated with the occurrence of high-grade lesions (univariate and multivariate analyses)

	Univariate analysis			Multivariate analysis		
	HR	IC95%	p	HR	IC95%	p
Female status	**2.6**	**0.9-6.4**	**0.0046**	1.6	0.4-6.0	0.59
Previously treated HR-AIN	**4.2**	**1.4-10.3**	**0.0008**	3.9	0.9-15.0	0.05
Age>44 years	**2.0**	**0.8-6.1**	**0.14**	1.4	0.5-4.4	0.53
HIV-positive status	0.8	0.3-2.4	0.66			
Tobacco consumption	1.1	0.3-4.1	0.75			
Obesity	0.9	0.1-3.1	0.89			
Previous HPV disease	1.6	0.8-3.4	0.21			
Anal sex	1.2	0.5-3.0	0.70			
Positive HPV16 at baseline	**3.0**	**1.3-6.4**	**0.0005**	**2.9**	**1.2-6.4**	**0.02**
Abnormal cytology at 1 year	**3.0**	**1.4-6.4**	**0.01**	2.4	1.0-5.3	0.08
HPV 16 & abnormal cytology	6.7	2.0-19.8	0.0001			