

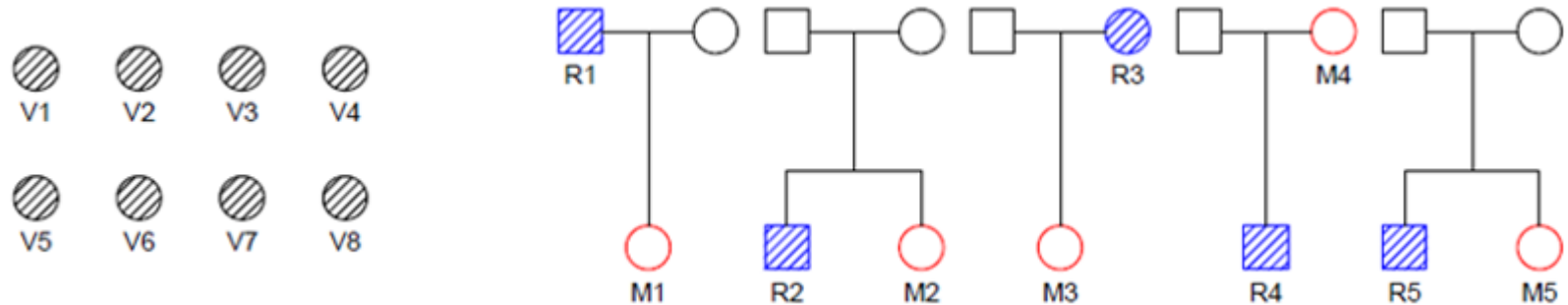


+ **dvir**

ISFG Summer School 2023 – Workshop 4.2



Lecture 6. Disaster victim identification



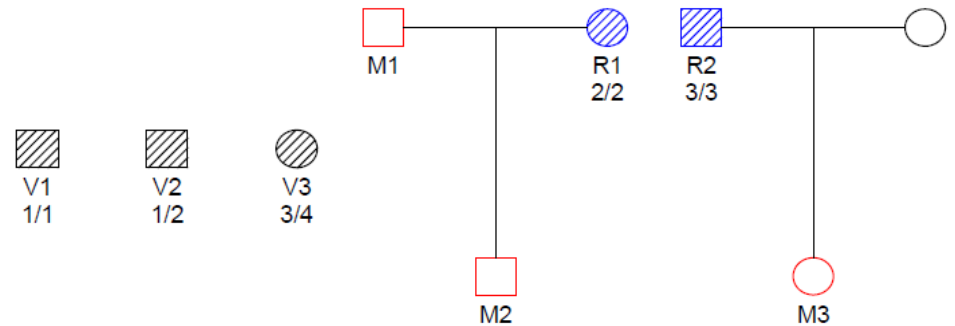
Thore Egeland

Norwegian University of Life Sciences (NMBU) &
Department of Forensic Medicine, Norway

Disaster Victim Identification (DVI)

- DVI

- Match list of unidentified persons against a list of missing persons
- Special case:
 - Missing Person Identification



- Cases

- Missing grandchildren of Argentina
- World Trade Center attack
- Spitsbergen civil aircraft disaster
- Balkan conflicts
- Drowned immigrants
- Thailand tsunami

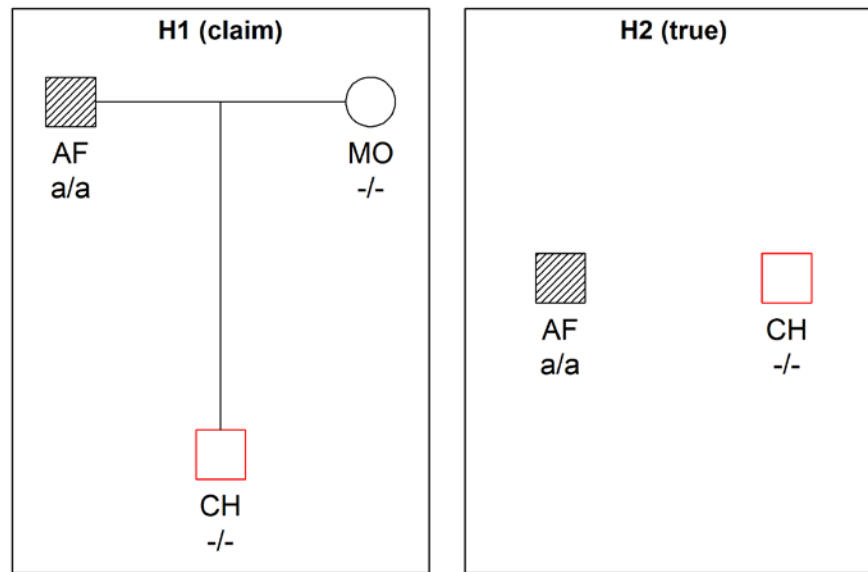
- Traditional methods and software

- Exclusion and inclusion
- Treat *each* victim (PM approach) or *each* family (AM approach) at a time
- Manual sequential approach
- Joint approach
 - library(dvir)



Exclusion power

- The *exclusion power* (EP) of a kinship test is the probability that H_1 ('claim') can be excluded, given that H_2 is true



$$\begin{aligned} EP &= P(\text{data incompatible with } H_1 \mid H_2) \\ &= P(\text{CH does not have } a \mid H_2) \\ &= (1 - p_a)^2 = (1 - 0.1)^2 = 0.81 \end{aligned}$$

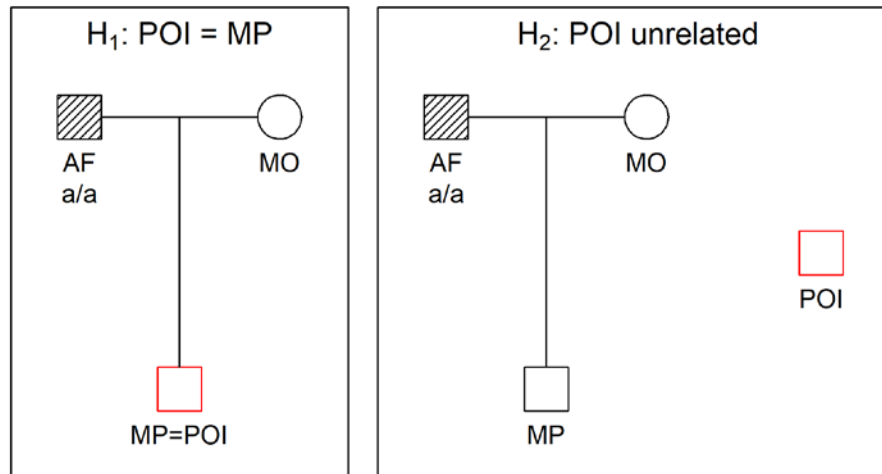
Exclusion power with the pedsuite

- The general function is
 - `exclusionPower(claimPed, truePed, ids)`
- If H_2 (true) is ‘unrelated’, we can use the simpler
 - `randomPersonEP(claimPed, id)`

```
> afr = c(a = 0.1, b = 0.9)
> nuclearPed(fa = "AF", child = "CH") |>
  addMarker(AF = "a/a", afreq = afr) |>
  randomPersonEP("CH")
```

```
Potential mismatches: 1 (1)
Expected mismatches: 0.81
P(at least 1 mismatch): 0.81
```

Inclusion power



- Do we enough data to give a positive match if POI = MP?
- **Inclusion power** (or *exceedance probability*):

$$IP_{10000} = P(LR > 10000 \mid POI = MP)$$

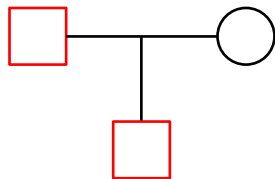
- Can be computed by simulation!
 - Unconditional → average for all pedigrees of this type
 - Conditional → probability for this particular case

In forrel:
> missingPersonIP()

Unconditional simulation

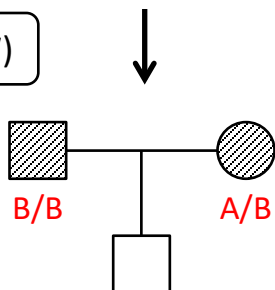
SNP: $p_A = p_B = 0.5$

A/A	A/B	B/B
0.25	0.5	0.25

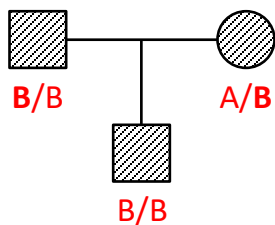


A/A	A/B	B/B
0.25	0.5	0.25

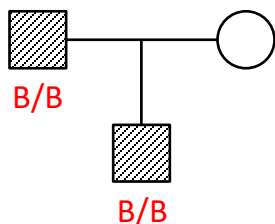
1. Founders (HW)



2. Gene dropping

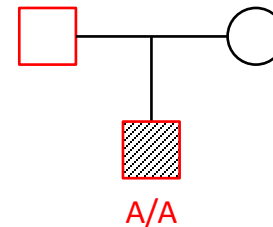


3. Remove untyped



Conditional simulation

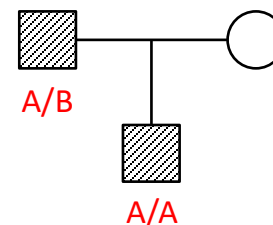
SNP: $p_A = p_B = 0.5$



1. Compute conditional distribution in the father

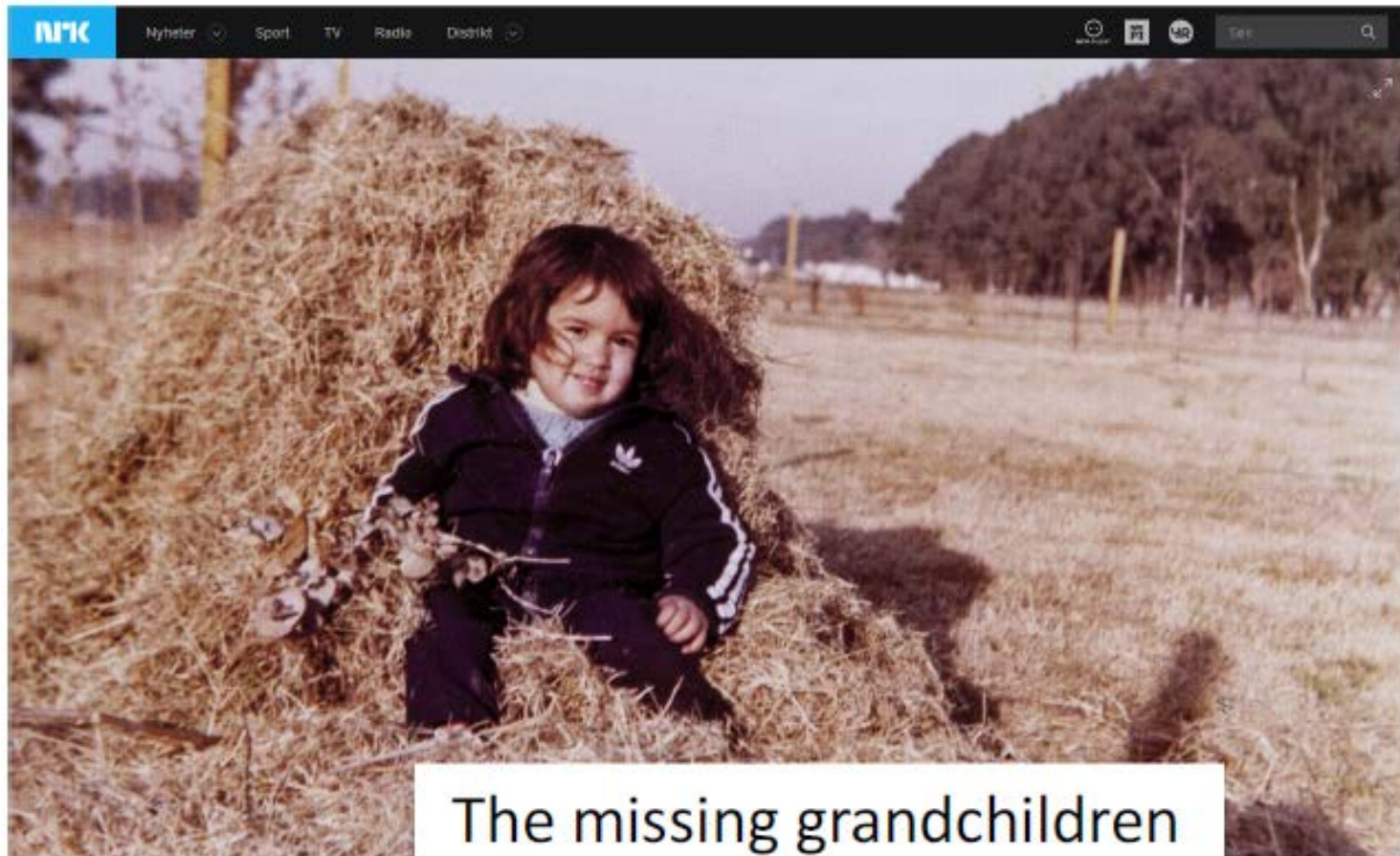
A/A	A/B	B/B
0.5	0.5	0

2. Sample from this



Computer intensive

Special case: missing person identification



The missing grandchildren
of Argentina

Argentina 1976 – 1983: Dirty War

- Military dictatorship
- *War* against guerrillas

- Opponents killed or disappeared

- *About 500 children abducted*
 - kidnapped with their parents or born in captivity
 - parents killed
 - raised by police or military families

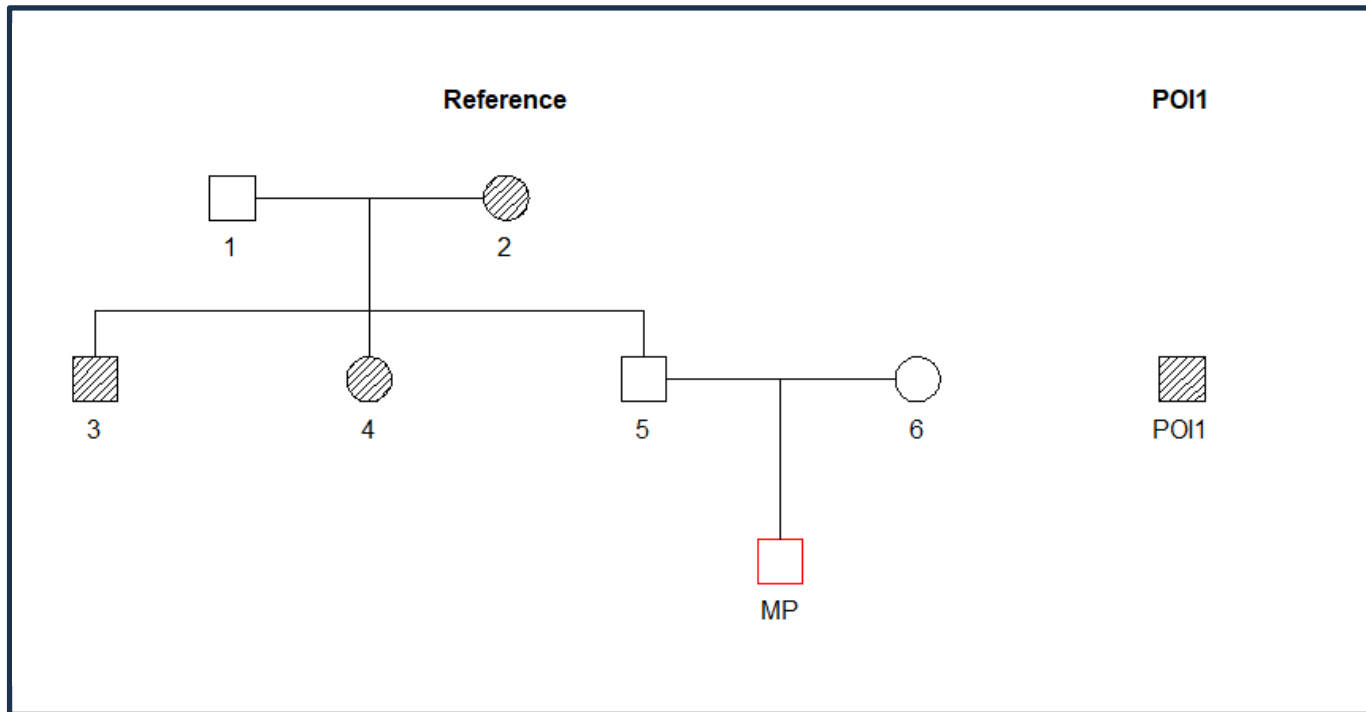


The missing grandchildren

- 1977: *Grandmothers of Plaza de Mayo*
 - weekly marches ever since
 - <https://www.abuelas.org.ar>
- 1983: Mary-Claire King, US geneticist
 - Collaboration with the Grandmothers
 - HLA typing + blood groups
 - First grandchild recovered using genetics
- 1987: National Genetic Data Bank (BNDG)
- 2007: Argentine Forensic Anthropology Team
- 2021: In total 130 reunifications by BNDG



Data



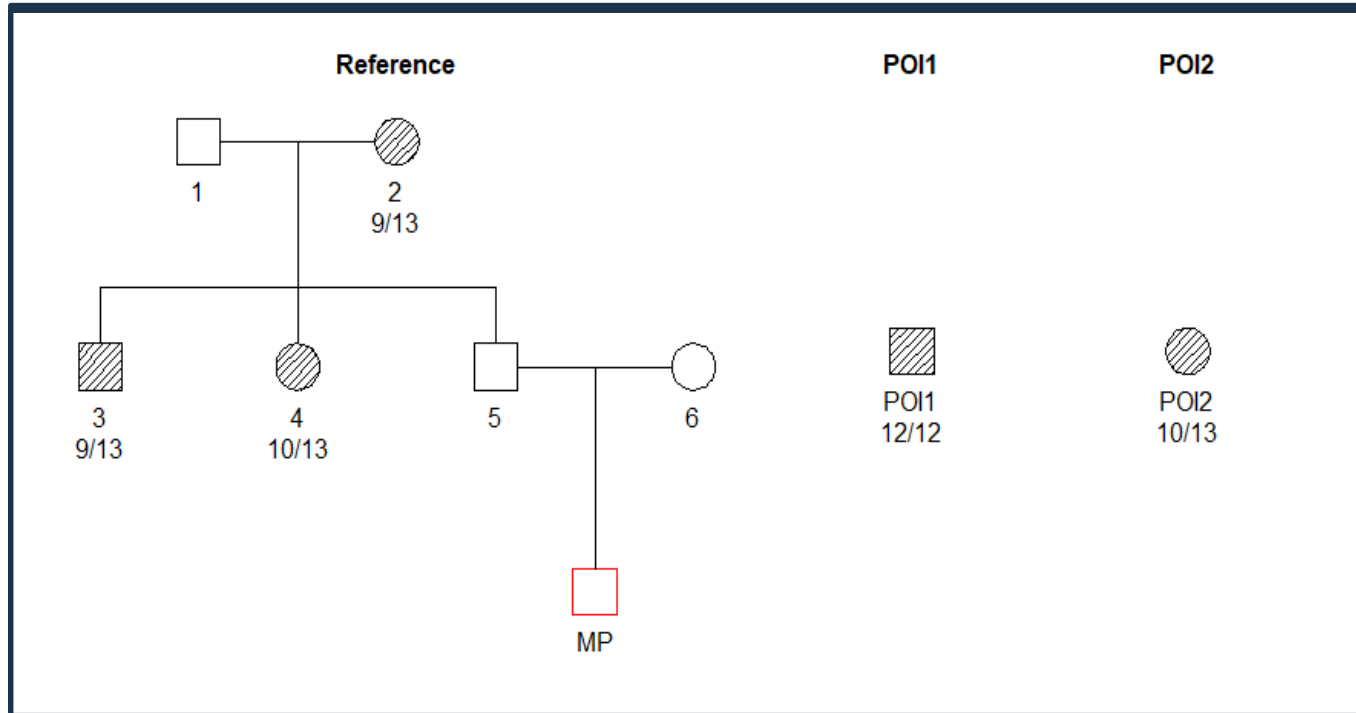
Standard marker kits:

- 15 - 24 autosomal STRs
- unlinked (mostly)

Sometimes also:

- X, Y, mtDNA markers
- Sex
- Age

Exclusion and inclusion



- POI1 = 12/12 is **excluded** if mutations are ignored since the alleles of his father are among 9/10/13
- POI2 is
 - ✓ **not excluded**
 - ✓ **included** if $LR > 10,000$ say.

Tailor made functions



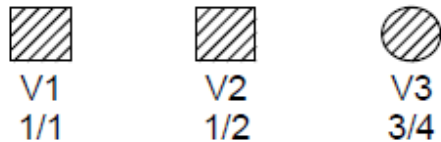
- `missingPersonPlot()`
- `missingPersonLR()`
- `missingPersonEP()`
- `missingPersonIP()`

- Exemplified in exercise

Pairwise approach

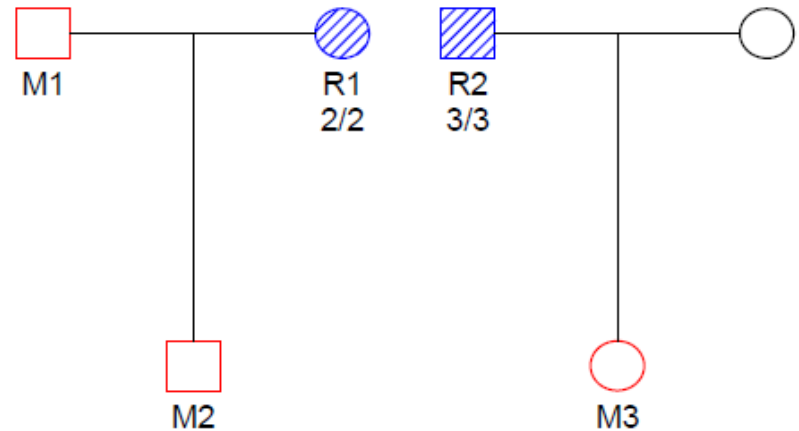
- **PM data**

- Victims V1, V2, ..., Vs
- Replicated samples merged



- **AM data**

- Missing persons M1, M2, ..., Mm
- References R1, R2, ...



- $p_1 = \dots = p_{10} = 0.1$
- Compare $V_2 = M_2$ to unrelated:

$$LR_{2,2} = \frac{p_2^2 \cdot p_1}{p_2^2 \cdot 2p_1p_2} = \frac{1}{2p_2} = 5.$$

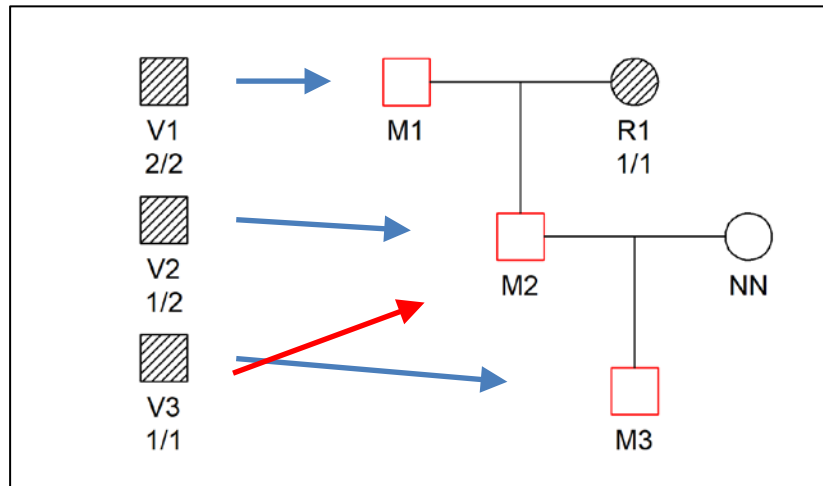
- All pairwise comparisons:

	M1	M2	M3
V1	1	0	0
V2	1	5	0
V3	0	0	5

Challenge: Sometimes pairwise fails

Joint solution

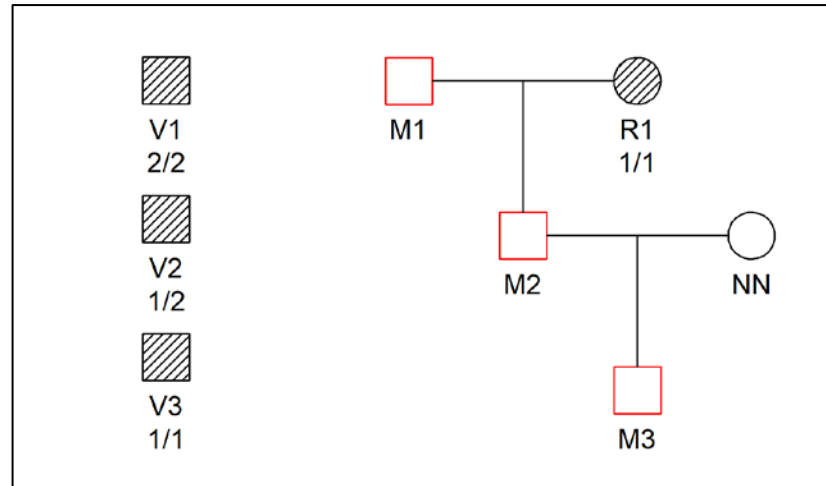
$V_1 = M_1$
 $V_2 = M_2$
 $V_3 = M_3$



Pairwise LRs

	M ₁	M ₂	M ₃
V ₁	1	0	0.5
V ₂	1	10	5.5
V ₃	1	20	10.5

Possible solutions: assignments



- **Assignment a**

- {V1 = M1, V2 = M2}
- compact notation: (M1, M2, *)
- 34 apriori possible assignments:

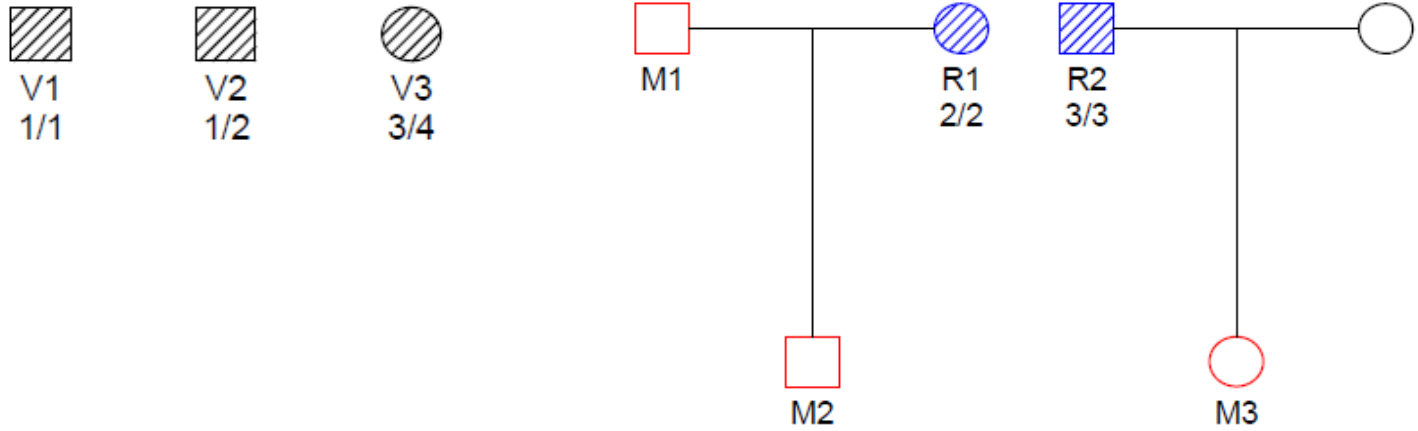
	V1	V2	V3
1	M3	M1	M2
2	M2	*	M1
3	M2	*	M3
4	*	M1	M2
...			
34	*	M2	*

The number of possible assignments

m_F	s_F					
	0	1	2	3	4	5
0	19081	3393	529	73	9	1
1	9276	3922	1074	228	40	6
2	4051	3135	1603	559	147	31
3	1546	2004	1768	1054	438	136
4	501	1045	1533	1533	1045	501
5	136	438	1054	1768	2004	1546
6	31	147	559	1603	3135	4051
7	6	40	228	1074	3922	9276
8	1	9	73	529	3393	19081

5 victims and 8 MPs

Joint approach: example



The model

- Assignments A
- Prior $\pi(a), a \in A$
- Posterior $\pi(a | \text{PM and AM data})$
- Likelihood

$$L(a) = P(\text{PM and AM data} \mid a, \Phi)$$

- Frequentist:

- $LR = L(a)/L_0$

- $L_0 = \text{likelihood empty assignment}$

- From this, we get

- $P(V_i = M_j \mid \text{PM and AM data})$

- $P(V_i = * \mid \text{PM and AM data})$

Example

	V_1	V_2	V_3	loglik	LR	posterior
1	M_1	M_2	M_3	-16.12	250.00	0.72
2	M_1	M_2	*	-17.73	50.00	0.14
3	*	M_2	M_3	-18.42	25.00	0.07
4	M_1	*	M_3	-20.03	5.00	0.01
5	*	M_1	M_3	-20.03	5.00	0.01
6	*	M_2	*	-20.03	5.00	0.01
7	*	*	M_3	-20.03	5.00	0.01
8	M_1	*	*	-21.64	1.00	0.00
9	*	M_1	*	-21.64	1.00	0.00
10	*	*	*	-21.64	1.00	0.00
11	M_2	M_1	M_3	-Inf	0.00	0.00
12	M_2	M_1	*	-Inf	0.00	0.00
13	M_2	*	M_3	-Inf	0.00	0.00
14	M_2	*	*	-Inf	0.00	0.00

	M_1	M_2	M_3	*
V_1	0.88	0.00	0.00	0.12
V_2	0.02	0.95	0.00	0.03
V_3	0.00	0.00	0.83	0.17

Numerical example

- The posterior with flat prior for the two first assignments:

$$\frac{250}{250 + 50 + \dots + 0} = 0.718$$

$$\frac{50}{250 + 50 + \dots + 0} = 0.144$$

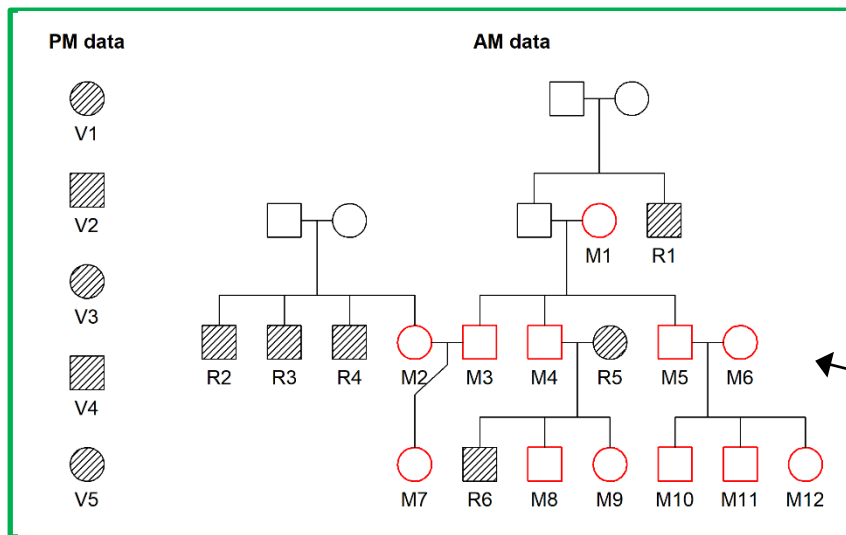
- Posterior pairing probability.
Sum rows where $V_1 = M_1$:

$$P(V_1 = M_1 \mid \text{PM and AM data}) = 0.718 + 0.144 + 0.0144 + 0.003 = 0.88$$

	V_1	V_2	V_3	loglik	LR	posterior
1	M_1	M_2	M_3	-16.118	250.000	0.718
2	M_1	M_2	*	-17.728	50.000	0.144
3	*	M_2	M_3	-18.421	25.000	0.072
4	M_1	*	M_3	-20.030	5.000	0.014
5	*	M_1	M_3	-20.030	5.000	0.014
6	*	M_2	*	-20.030	5.000	0.014
7	*	*	M_3	-20.030	5.000	0.014
8	M_1	*	*	-21.640	1.000	0.003
9	*	M_1	*	-21.640	1.000	0.003
10	*	*	*	-21.640	1.000	0.003

	M_1	M_2	M_3	*
V_1	0.88	0.00	0.00	0.12
V_2	0.02	0.95	0.00	0.03
V_3	0.00	0.00	0.83	0.17

Large example (ICMP 2017)



Sorted assignments: 5 most likely solutions

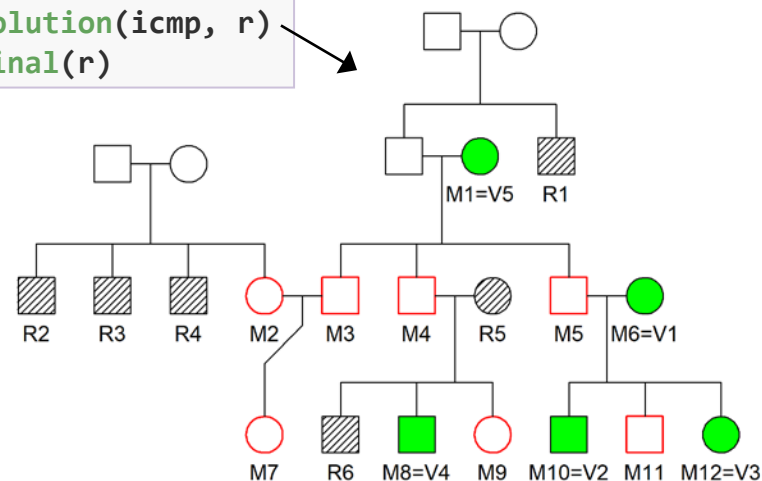
	V ₁	V ₂	V ₃	V ₄	V ₅	loglik	LR	posterior
1	M ₆	M ₁₀	M ₁₂	M ₈	M ₁	-312.98	1.14E+24	0.50
2	M ₆	M ₁₁	M ₁₂	M ₈	M ₁	-312.98	1.14E+24	0.50
3	M ₆	M ₁₀	M ₁₂	M ₈	M ₇	-327.16	7.86E+17	0.00
4	M ₆	M ₁₁	M ₁₂	M ₈	M ₇	-327.16	7.86E+17	0.00
5	M ₆	*	M ₁₂	M ₈	M ₁	-327.74	4.40E+17	0.00

R code

```
library(dvir)
plotDVI(icmp)
r = jointDVI(icmp)
plotSolution(icmp, r)
Bmarginal(r)
```

Posterior pairing probabilities

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	*
V1						1							
V2										0.5	0.5		
V3													1
V4								1					
V5	1												



Disaster Victim Identification in R: dvir package

- Key functions in exercise:

- `plotDVI()`

- `pairwiseLR()`

- `findExcluded()`

- `findUndisputed`

- `jointDVI()`

- `plotSolution()`

```
> excl$exclusionMatrix
  MP1 MP2 MP3 MP4 MP5 MP6 MP7 MP8
V1  0 NA  4  2  2  9 NA NA
V2 NA  0 NA NA NA NA  4  0
V3  4 NA  0  2  2  9 NA NA
V4  6 NA  4  0  0  5 NA NA
V5  6 NA  4  0  0  7 NA NA
V6 10 NA  5  1  1  0 NA NA
V7 NA 12 NA NA NA NA  0  0
V8 NA  9 NA NA NA NA  2  0
```

```
4 undisputed matches
```

```
V1 = MP1 (LR = 4.8e+08)
```

```
V2 = MP2 (LR = 6.78e+10)
```

```
V3 = MP3 (LR = 6.41e+14)
```

```
V6 = MP6 (LR = 8.82e+12)
```

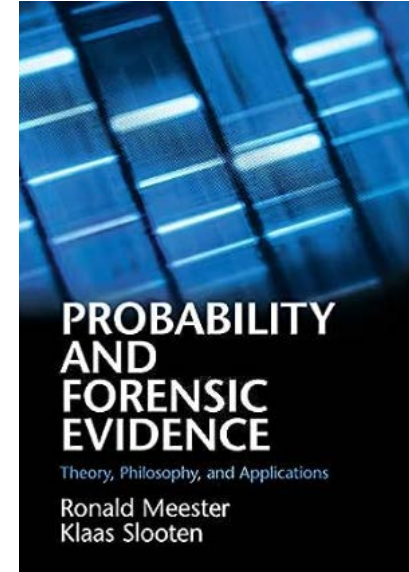
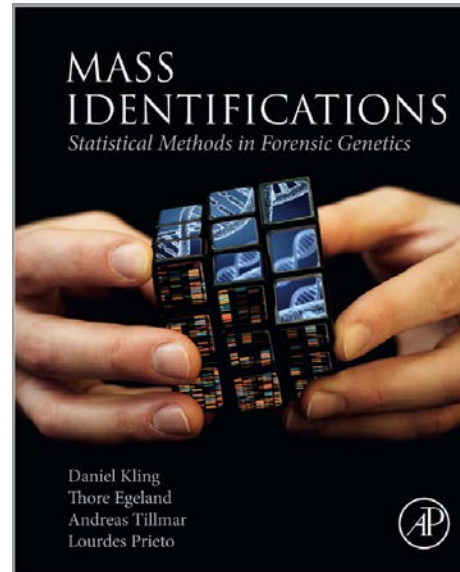
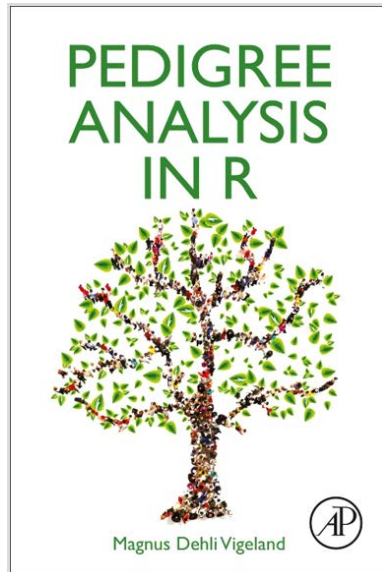
Discussion

- What should be reported?
 - LR and, or posterior?
 - Joint or individual?
 - Use verbal scales?
- Extend model
 - informative priors
 - include non-genetic information in likelihood

$$\begin{aligned}L(a) &= P(PM, AM, age|a, \Phi) \\ &= P(PM, AM|a, \Phi)P(age|a, \Phi)\end{aligned}$$

- beyond conventional forensic markers

Biased selection of references



Article | [Open Access](#) | Published: 01 July 2021

Joint DNA-based disaster victim identification

Magnus D. Vigeland [✉](#) & Thore Egeland

Scientific Reports **11**, Article number: 13661 (2021) | [Cite this article](#)