



Istituto Nazionale di Fisica Nucleare
Sezione di Napoli



HIGH ENERGY PHYSICS MEETS NUTRITION SCIENCE

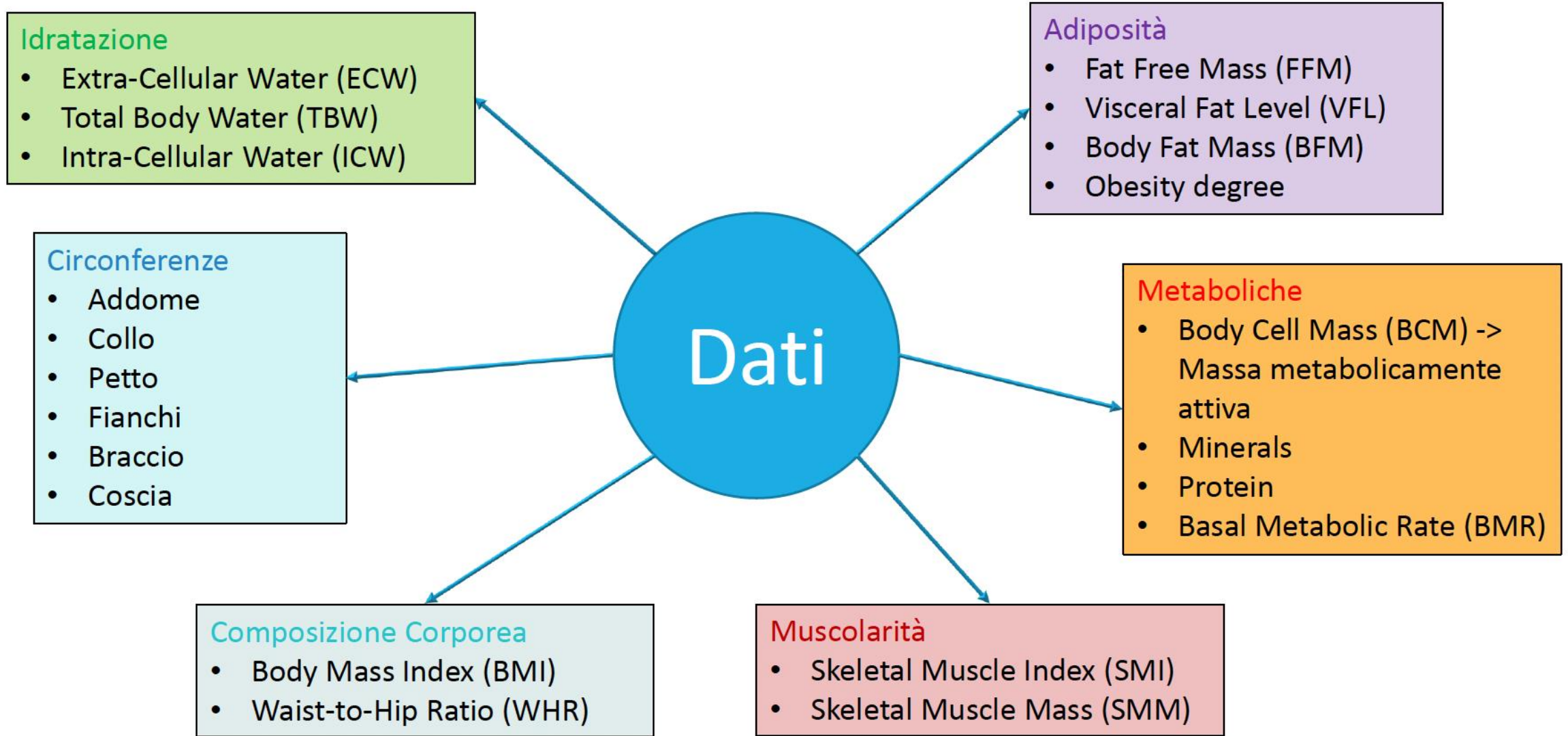
Machine Learning 4 Nutrition Science Project (ML4N):
II Workshop

Statistical description of the patients' sample

Francesco Cirotto, Francesco Conventi, Antonio D'Avanzo, Agostino De Iorio,
Giovanni Perna, Biagio Rossi, Elvira Rossi

Istituto Nazionale Fisica Nucleare
Università degli studi di Napoli "Federico II" e "Parthenope"

Variabili disponibili



Suddivisione in cluster

	Uomini	Donne	Totale
Pazienti	424	1202	1626
Totale visite	1436	5520	6664

Il campione è stato **separato in cluster** per individuare sottogruppi con caratteristiche comuni.

In totale sono stati individuati **10 cluster**:

- **4** esclusivamente o in larga maggioranza **composti da uomini**
- **6** in larga maggioranza composti **da donne**

Scopo dell'analisi

Studiare gli andamenti temporali delle variabili direttamente connesse allo stato di salute e benessere del paziente

Requisiti

- I. Pazienti con almeno due visite
- II. Eliminare eventuali errori nella registrazione dei dati

Il campione studiato nel dettaglio è composto da:

	Uomini	Donne	Totale
Pazienti	424	1202	1626
Totale visite	1436	5520	6664

Considerazioni generali - Uomini

Valore medio per il cluster:	1	2	3	4	5	6	7	8	9	10
Sesso	0.00	0.89	0.99	0.62	0.21	0.90	0.94	0.04	0.08	0.65
Età	41.50	44.39	55.00	39.16	38.52	27.11	34.57	38.09	36.06	22.90
Altezza [cm]	182.00	160.98	157.63	164.00	173.83	161.23	161.49	179.12	178.45	152.26
Peso [kg]	184.50	100.60	76.14	129.14	86.60	87.31	67.96	133.15	109.92	43.54
BMI [kg/ m ²]	55.80	38.83	30.66	48.07	28.72	33.62	26.04	41.45	34.56	18.82
BFM [%]	50.30	49.15	43.38	52.20	28.63	45.51	34.44	42.26	36.76	20.54
FFM [%]	49.70	50.85	56.62	47.80	71.37	54.49	65.56	57.74	63.24	79.46
Numero totale di pazienti	4	136	224	16	300	107	707	24	73	34

- Cluster 1 costituito da 4 soggetti affetti da **obesità grave**
- Cluster 8 costituito da 24 soggetti molto sovrappeso
- Cluster 9 costituito da 73 soggetti **sovrappeso**
- Cluster 5 costituito da 300 soggetti **leggermente sovrappeso**

Considerazioni generali - Donne

Valore medio per il cluster:	1	2	3	4	5	6	7	8	9	10
Sesso	0.00	0.89	0.99	0.62	0.21	0.90	0.94	0.04	0.08	0.65
Età	41.50	44.39	55.00	39.16	38.52	27.11	34.57	38.09	36.06	22.90
Altezza [cm]	182.00	160.98	157.63	164.00	173.83	161.23	161.49	179.12	178.45	152.26
Peso [kg]	184.50	100.60	76.14	129.14	86.60	87.31	67.96	133.15	109.92	43.54
BMI [kg/ m ²]	55.80	38.83	30.66	48.07	28.72	33.62	26.04	41.45	34.56	18.82
BFM [%]	50.30	49.15	43.38	52.20	28.63	45.51	34.44	42.26	36.76	20.54
FFM [%]	49.70	50.85	56.62	47.80	71.37	54.49	65.56	57.74	63.24	79.46
Numero totale di pazienti	4	136	224	16	300	107	707	24	73	34

- **Cluster 4** costituito da 16 soggetti affetti da **obesità grave di età media di 39 anni**
- **Cluster 2** costituito da 136 soggetti **sovrappeso e meno giovani**
- **Cluster 6** costituito da 107 soggetti **sovrappeso e molto giovani**
- **Cluster 7** costituito da 707 soggetti **leggermente sovrappeso e giovani**
- **Cluster 3** costituito da 224 soggetti **leggermente sovrappeso e meno giovani**
- **Cluster 10** costituito da 34 soggetti **sottopeso e molto sottopeso e molto giovani**

Distribuzioni della popolazione: WHR

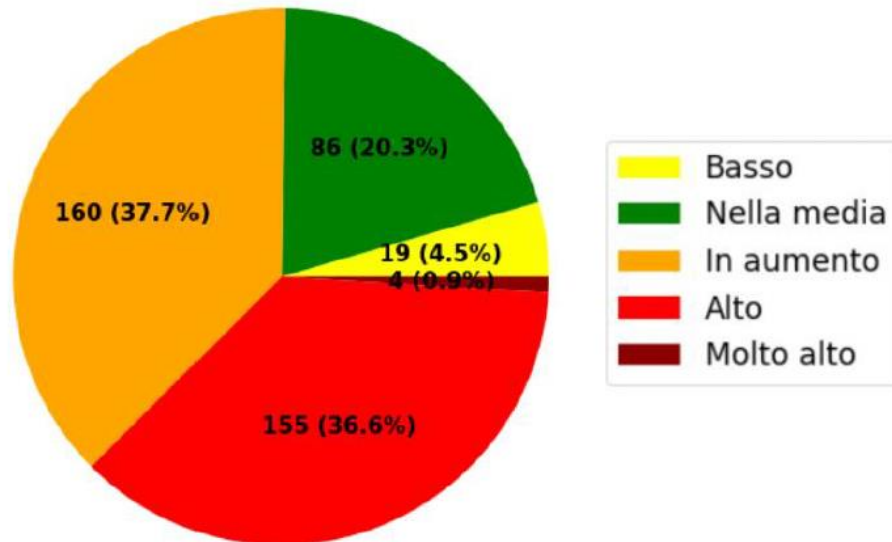
WHR

La variabile **WHR** è il rapporto adimensionale tra la circonferenza della vita e quella dei fianchi.

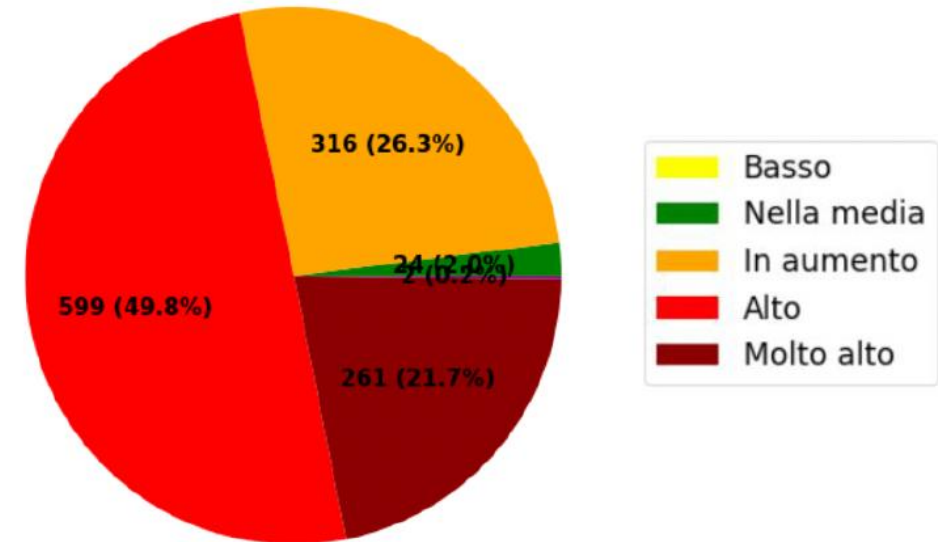
È associato al **rischio di malattia cardiovascolare**

Rischio di malattia	Uomini	Donne
Basso	< 0.80	< 0.70
Nella media	0.81 – 0.90	0.71 – 0.80
In aumento	0.91 – 0.99	0.81 – 0.89
Alto	1.00 – 1.19	0.90 – 1.09
Molto alto	1.20 – 1.29	1.10 – 1.19
Estremamente alto	> 1.30	> 1.20

WHR Uomini



WHR Donne



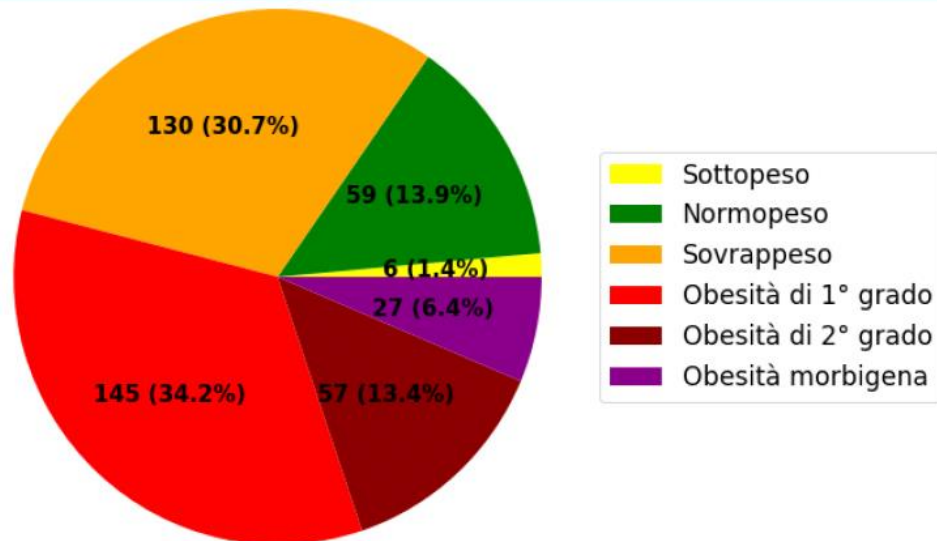
Distribuzioni della popolazione: Body Mass Index

BMI

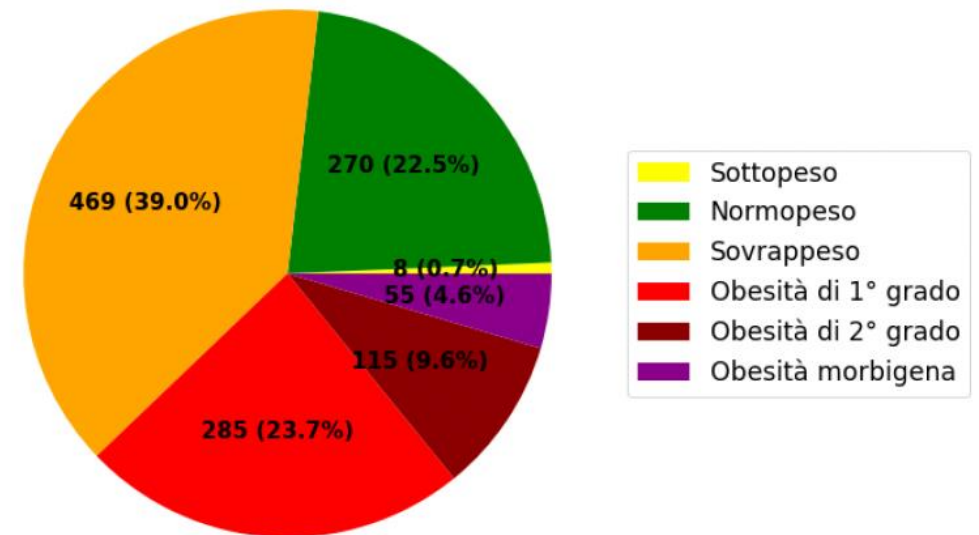
L'indice di massa corporea è un dato biometrico, espresso come rapporto tra peso e quadrato dell'altezza di un individuo ed è utilizzato come un indicatore dello stato di peso forma.

Valori di riferimento	
Sottopeso	≤ 18.5
Normopeso	18.6 - 24.9
Sovrappeso	25.0 - 29.9
Obesità 1° grado	30.0 - 34.9
Obesità 2° grado	35.0 - 39.9
Obesità morbigena	≥ 40

BMI Uomini



BMI Donne



Soggetti sottopeso

Valore medio per il cluster:	10
Sesso	0.65
Età	22.90
Altezza [cm]	152.26
Peso [kg]	43.54
BMI [kg/ m ²]	18.82
BFM [%]	20.54
FFM [%]	79.46
Circonferenza braccio [cm]	24.64
Circonferenza collo [cm]	29.24
Circonferenza torace [cm]	78.22
Circonferenza addome [cm]	66.56
Circonferenza fianchi [cm]	84.54
Circonferenza braccio destro [cm]	24.72
Circonferenza braccio sinistro [cm]	24.64
Circonferenza coscia destra [cm]	44.62
Circonferenza coscia sinistra [cm]	44.60
Numero totale di pazienti	34

Cluster 10:

- costituito da 34 soggetti **sottopeso**
- Soggetti mediamente molto giovani (23 anni)
- Proporzioni tra massa grassa e massa magra molto vicine ai valori normali:
 - BFM [14 - 19]
 - FFM [81 - 86]
- Circonferenze tutte più basse della media

Obiettivi della dieta:

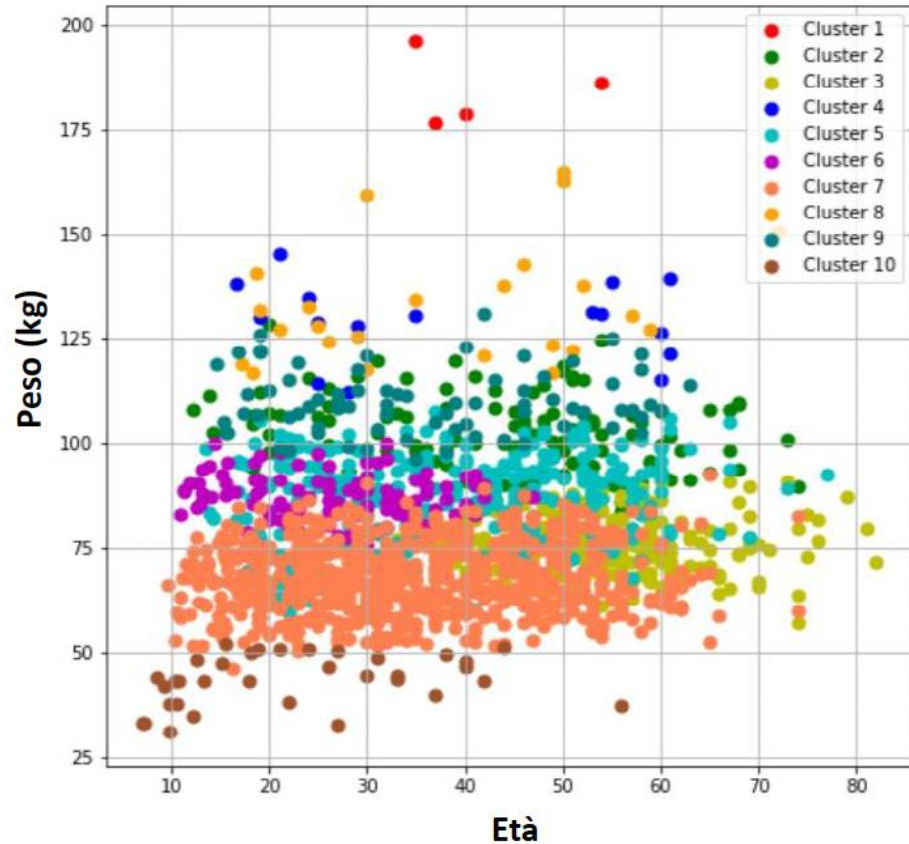
- Aumento di peso → aumento del BMI
- Normalizzazione delle proporzioni BFM e FFM

Variabili principali per la suddivisione

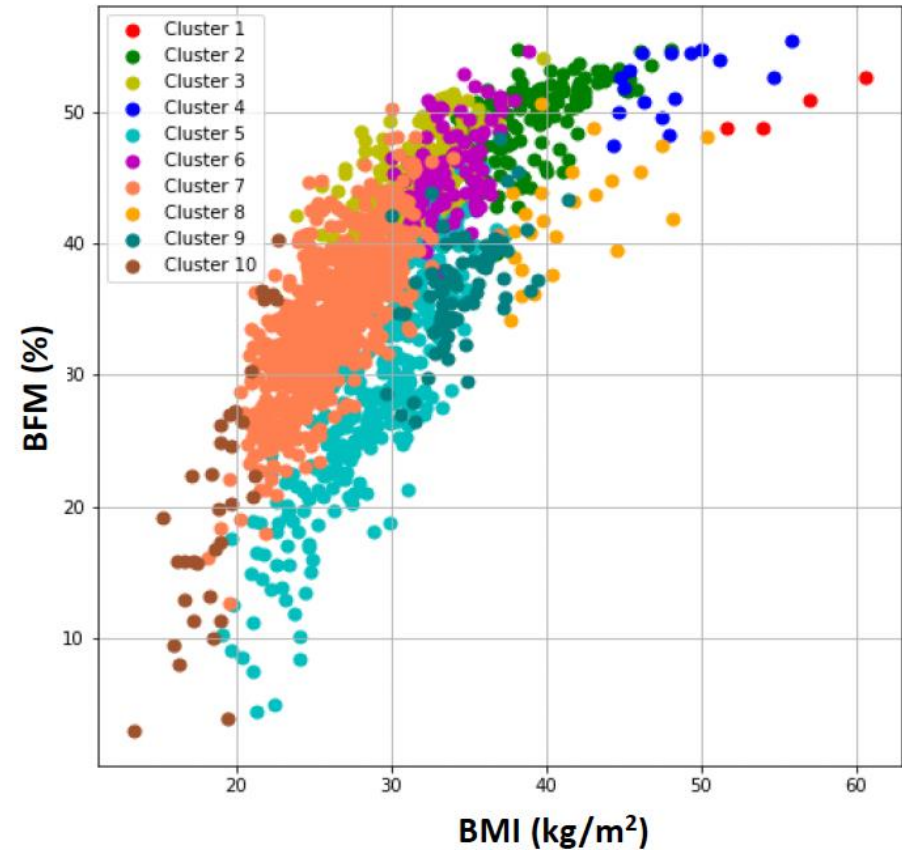
La tecnica di ML utilizzata ha trovato gruppi di individui con caratteristiche simili.

Ma quali sono le caratteristiche simili?

Età e peso



BMI e percentuale massa grassa (BFM)



Soggetti leggermente sovrappeso

Valore medio per il cluster:	5	7
Sesso	0.21	0.94
Età	38.52	34.57
Altezza [cm]	173.83	161.49
Peso [kg]	86.60	67.96
BMI [kg/·m ²]	28.72	26.04
BFM [%]	28.63	34.44
FFM [%]	71.37	65.56
Circonferenza braccio [cm]	35.20	31.69
Circonferenza collo [cm]	38.93	35.03
Circonferenza torace [cm]	104.99	93.92
Circonferenza addome [cm]	99.17	88.40
Circonferenza fianchi [cm]	104.86	97.86
Circonferenza braccio destro [cm]	35.29	31.74
Circonferenza braccio sinistro [cm]	35.20	31.69
Circonferenza coscia destra [cm]	57.45	52.99
Circonferenza coscia sinistra [cm]	57.12	52.77
Numero totale di pazienti	300	707

Cluster 5:

- Costituito da 300 soggetti
- Maggiormente uomini (79 %)
- Soggetti mediamente giovani (media 38 anni)

Cluster 7:

- Costituito da 707 soggetti
- Maggiormente donne (94 %)
- Soggetti mediamente giovani (media 34 anni)

- Proporzioni tra massa grassa e massa magra molto prossime ai valori di riferimento:
 - BFM [19 - 23]
 - FFM [77 - 81]
- Circonferenze della media

Obiettivi della dieta:

- Lieve perdita di peso → lieve abbassamento del BMI
- Mantenimento

Soggetti obesi

Valore medio per il cluster:	1	4
Sesso	0.00	0.62
Età	41.50	39.16
Altezza [cm]	182.00	164.00
Peso [kg]	184.50	129.14
BMI [kg/ m ²]	55.80	48.07
BFM [%]	50.30	52.20
FFM [%]	49.70	47.80
Circonferenza braccio [cm]	83.25	52.06
Circonferenza collo [cm]	50.30	50.59
Circonferenza torace [cm]	130.30	126.27
Circonferenza addome [cm]	153.53	134.27
Circonferenza fianchi [cm]	141.35	128.84
Circonferenza braccio destro [cm]	82.88	52.13
Circonferenza braccio sinistro [cm]	83.25	52.06
Circonferenza coscia destra [cm]	73.23	68.74
Circonferenza coscia sinistra [cm]	69.95	66.93
Numero totale di pazienti	4	16

Cluster 1:

- Esclusivamente uomini
- Soggetti mediamente giovani (40-50 anni)

Cluster 4:

- Maggioranza donne
- Due popolazioni:
 - molto giovane (25 anni)
 - meno giovane (55 anni)

- Proporzioni tra massa grassa e massa magra molto lontane da valori normali:
 - BFM [18 - 24]
 - FFM [76 - 82]
- Circonferenze tutte più alte della media

Obiettivi della dieta:

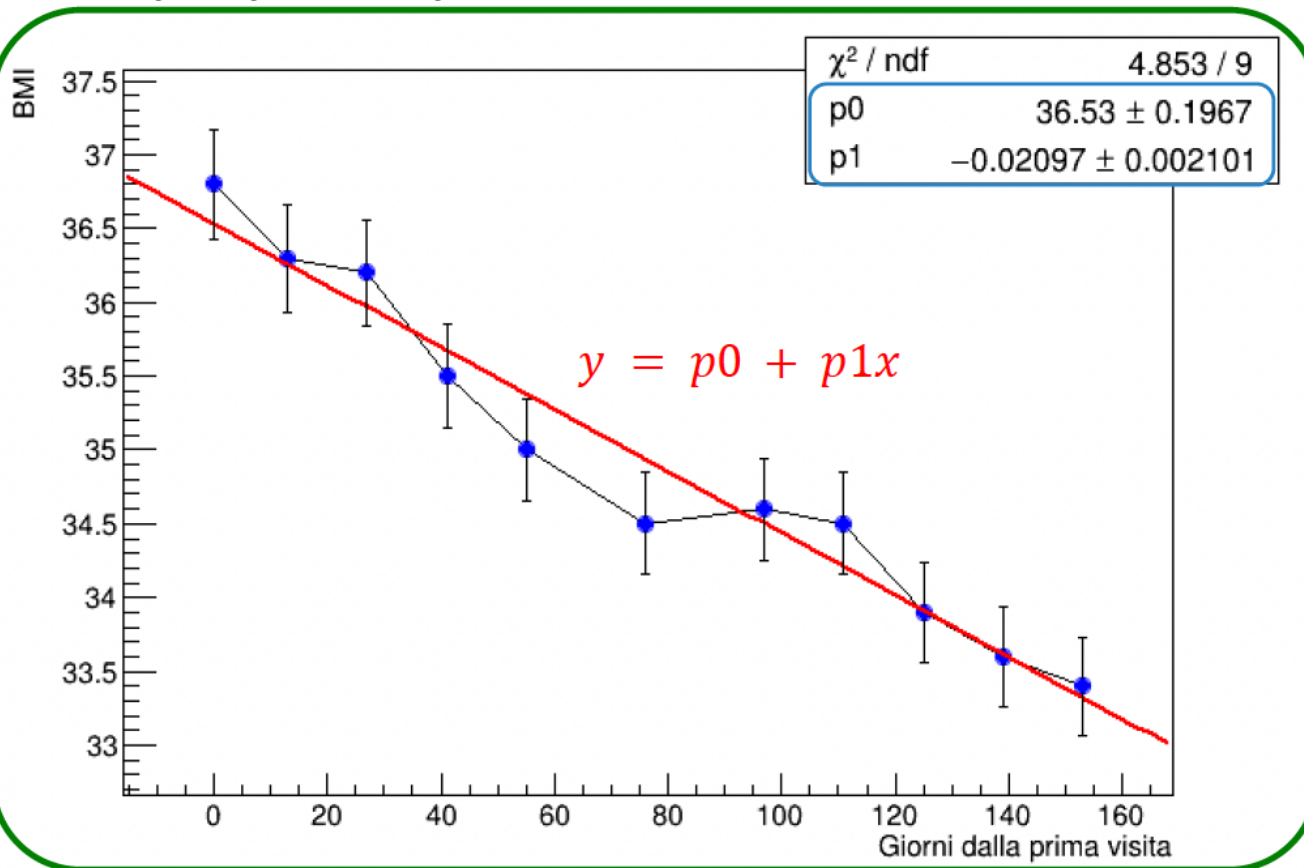
- Sensibile perdita di peso → abbassamento del BMI
- Normalizzazione delle proporzioni BFM e FFM
- Mantenimento della Massa Muscolare
- Miglioramento dell'idratazione

Andamento della dieta nel tempo

La nostra analisi si basa sullo studio di alcune variabili più indicative dello stato di salute del paziente.

L'obiettivo è monitorare i parametri nel tempo e studiarne l'andamento.

Esempio per un paziente:



Interpretazione

Riportiamo in un grafico i valori di una variabile (BMI) per ogni visita

Le visite sono riportate in giorni dopo la prima visita

Per l'andamento viene effettuato un **fit lineare**:
viene cioè adattata una retta

$$y = p_0 + p_1x$$

che rappresenta il modello con cui descriviamo i dati

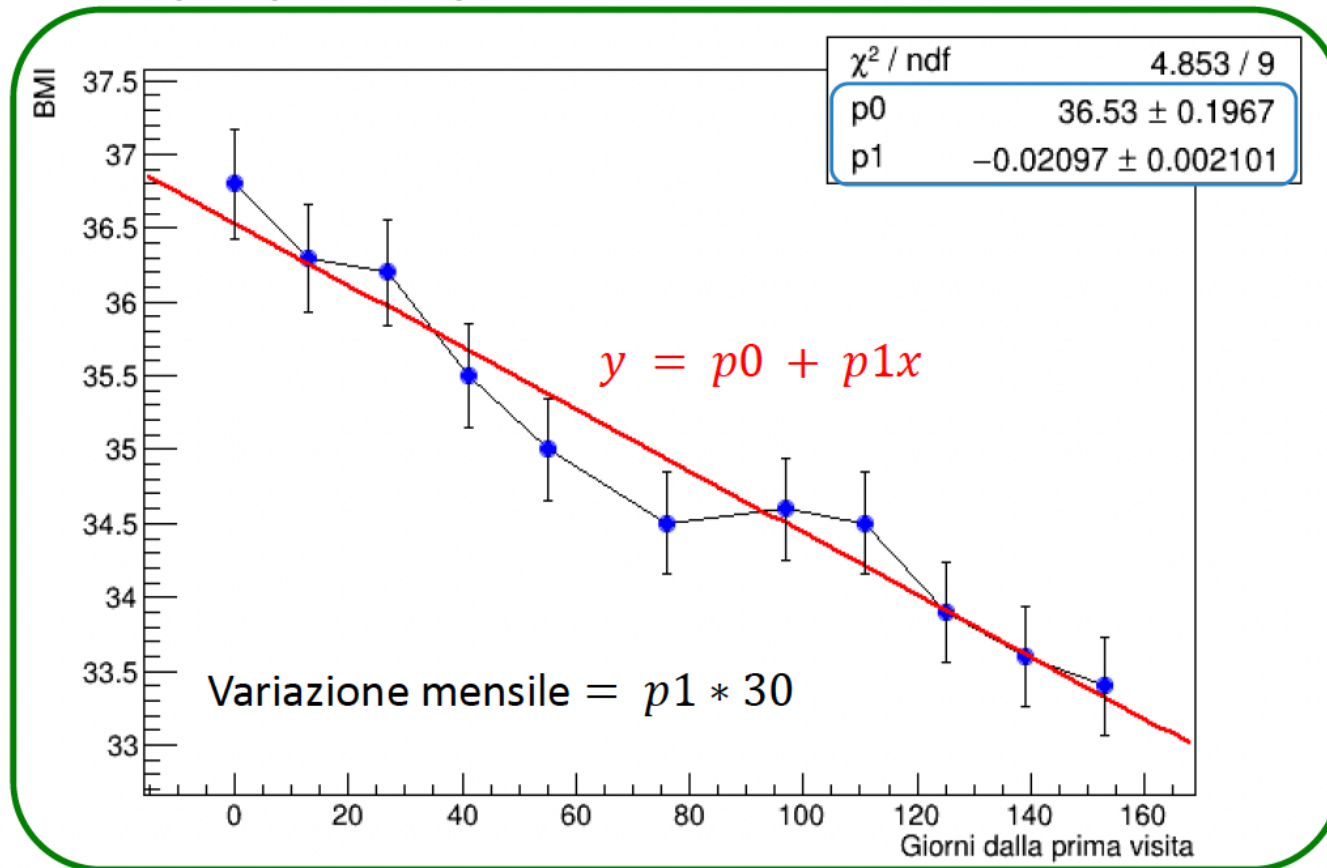
I parametri della retta p_0 e p_1 rappresentano il valore iniziale e la variazione giornaliera della variabile

Andamento della dieta nel tempo

La nostra analisi si basa sullo studio di alcune variabili più indicative dello stato di salute del paziente.

L'obiettivo è monitorare i parametri nel tempo e studiarne l'andamento.

Esempio per un paziente:



Procedimento

Il parametro p_1 viene moltiplicato per 30 al fine di ottenere la variazione mensile

Questo studio viene ripetuto per ogni paziente analizzato

Risultato

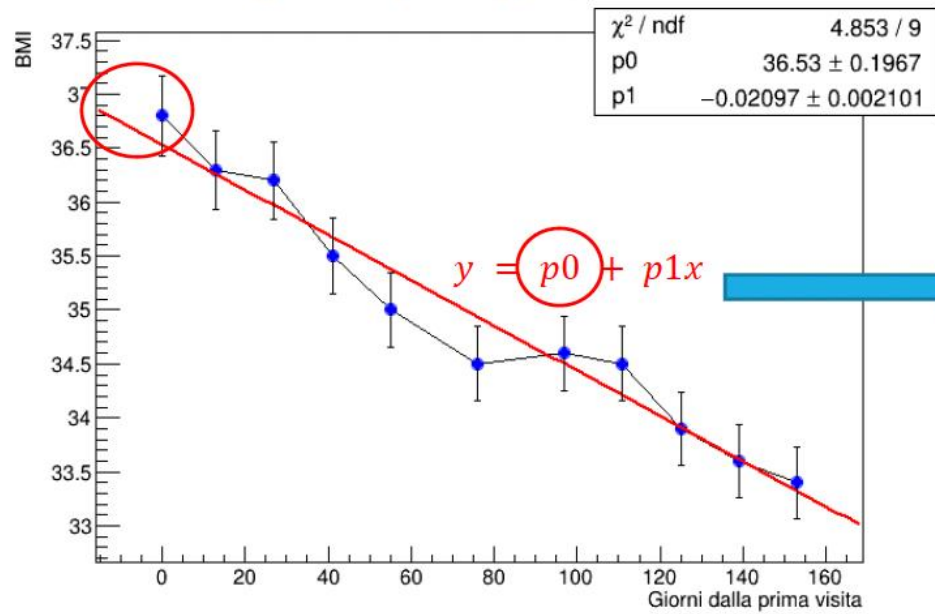
Collezione di andamenti nel tempo

Generalizzazione del valore iniziale

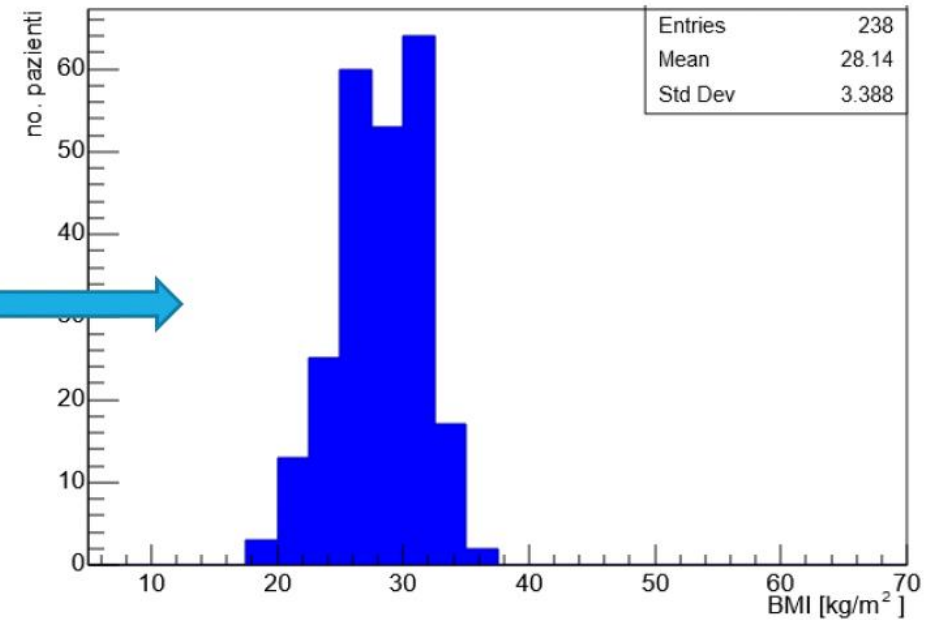
I valori ottenuti per il valore iniziale della variabile in esame, ripetuto per ogni paziente, viene riportato in un istogramma

Istogramma complessivo dei valori iniziali della variabile in esame per un singolo cluster.
Esempio per cluster 5

Ripetuto per ogni paziente

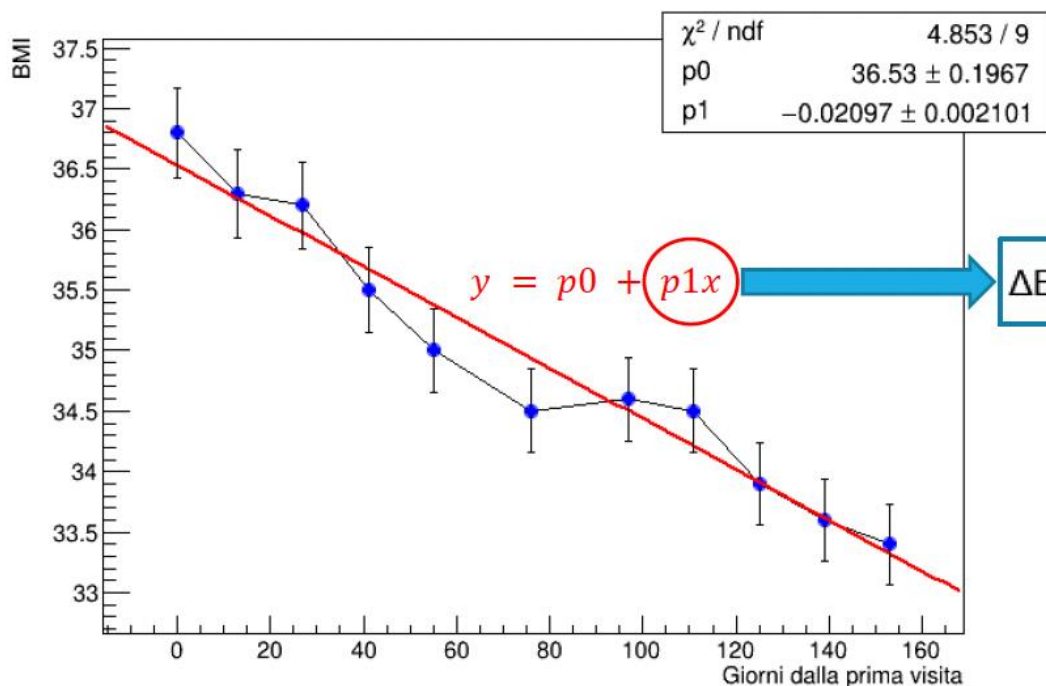


BMI iniziale = p_0

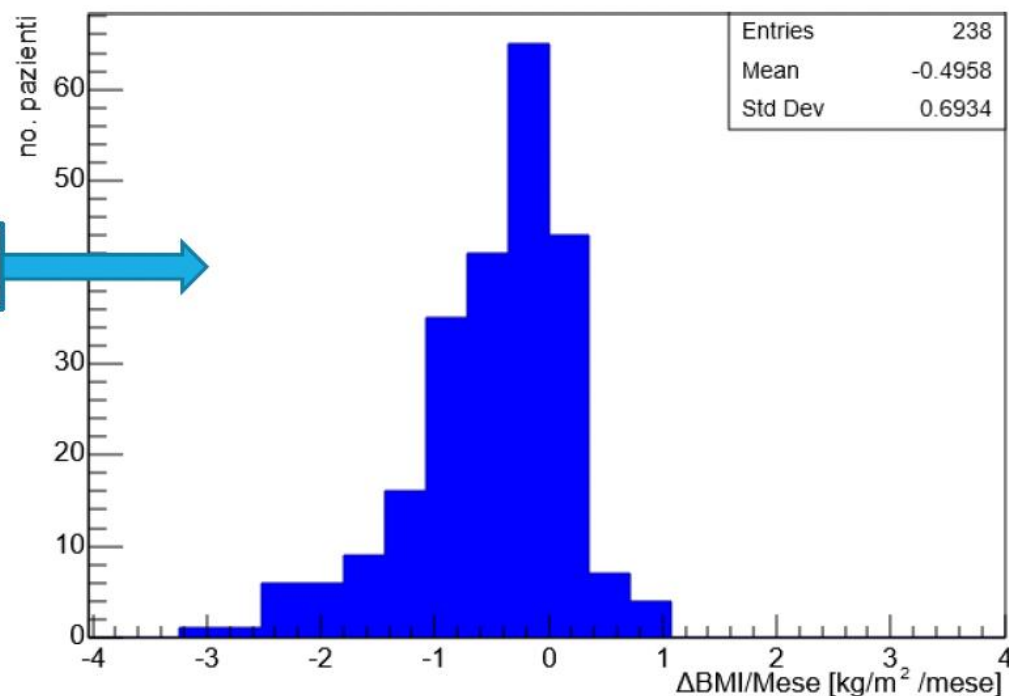


Generalizzazione della variazione mensile

I valori ottenuti per il parametro della variazione mensile, ripetuto per ogni paziente, viene riportato in un istogramma



Istogramma complessivo dei parametri di variazione mensile per un singolo cluster. Esempio per cluster 5



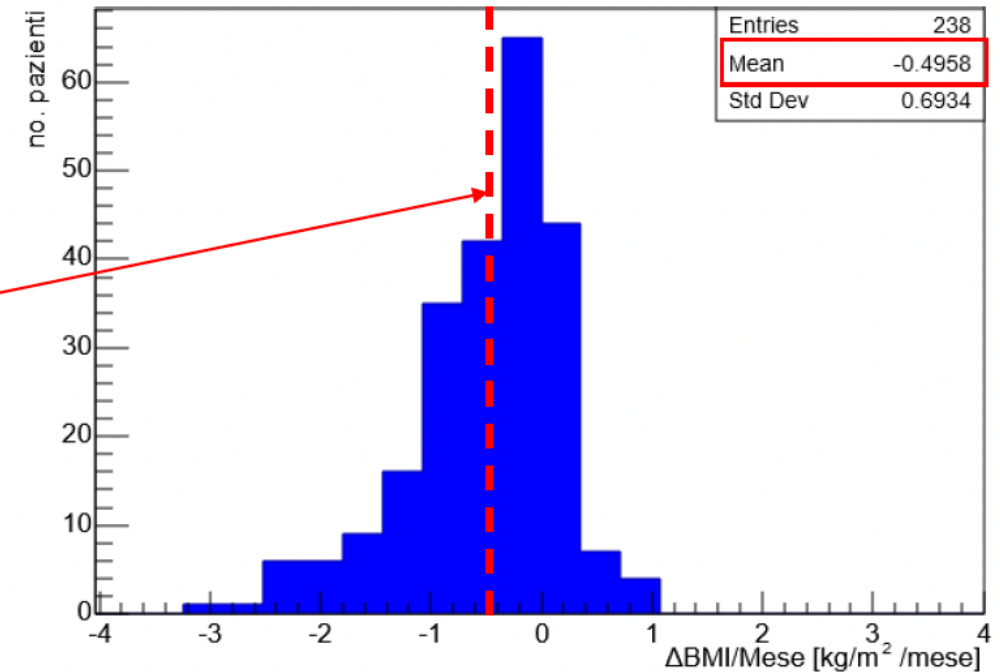
Interpretazione della distribuzione

Il **valore medio** di questo istogramma ci dice come si comporta il cluster:

Un valore medio negativo vuol dire che mediamente le persone appartenenti a quel cluster stanno diminuendo il BMI e dunque perdendo peso

Un valore medio positivo vuol dire che mediamente le persone appartenenti a quel cluster stanno aumentando il BMI e dunque aumentando peso

Istogramma complessivo dei parametri di variazione mensile per un singolo cluster.
Esempio per cluster 5



Interpretazione degli andamenti temporali

Esempio:

Gli uomini appartenenti al cluster 5 stanno perdendo circa mezzo punto di BMI al mese

Domanda

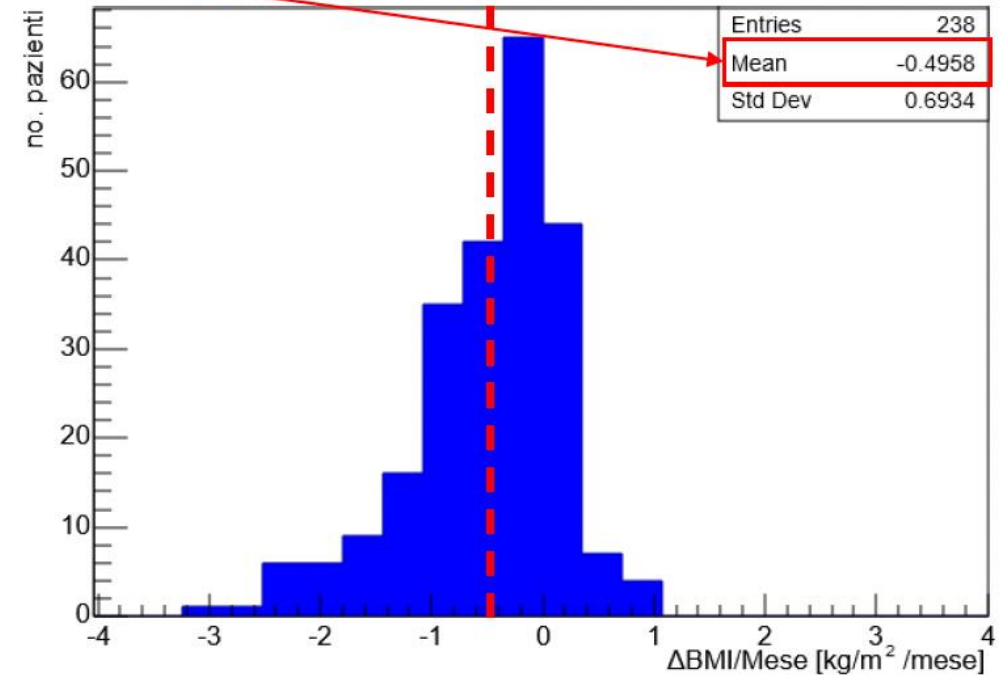
È un andamento positivo per la dieta?

È quello che ci aspettiamo?

Valore medio per il cluster: 5

Sesso	0.21
Età	38.52
Altezza [cm]	173.83
Peso [kg]	86.60
BMI [kg/cm ²]	28.72
BFM [%]	28.63
FFM [%]	71.37

Istogramma complessivo dei parametri di variazione mensile per un singolo cluster.
Esempio per cluster 5



Andamenti temporali e previsioni

Esempio:

Gli uomini appartenenti al cluster 5 stanno perdendo circa mezzo punto di BMI al mese

Domanda

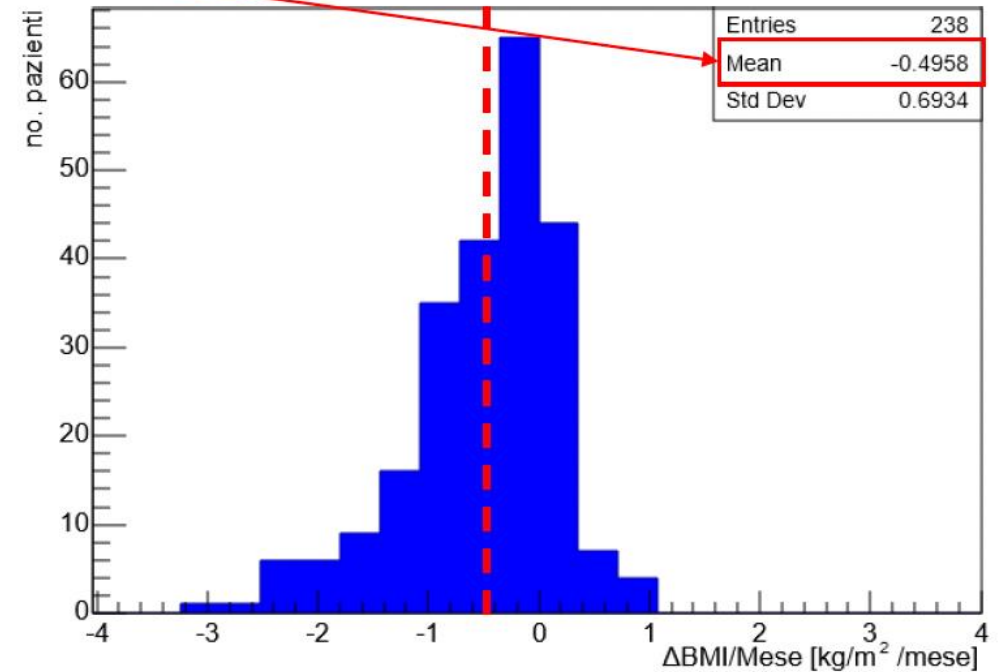
È un andamento positivo per la dieta?
È quello che ci aspettiamo?

Valore medio per il cluster: 5

Sesso	0.21
Età	38.52
Altezza [cm]	173.83
Peso [kg]	86.60
BMI [kg/cm ²]	28.72
BFM [%]	28.63
FFM [%]	71.37

Sì! Questo cluster parte da un valore di circa 29 e punta ad arrivare al valore di riferimento di 25

Istogramma complessivo dei parametri di variazione mensile per un singolo cluster.
Esempio per cluster 5



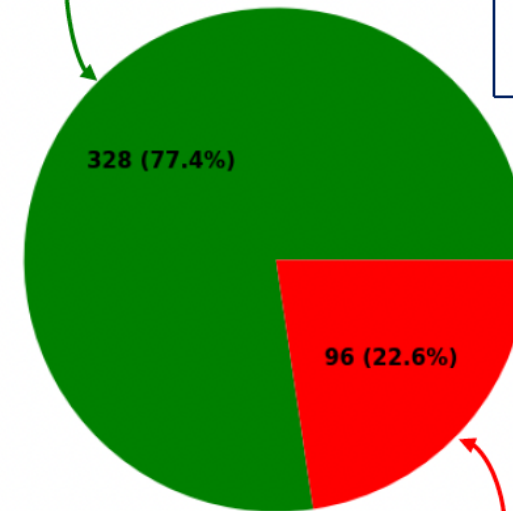
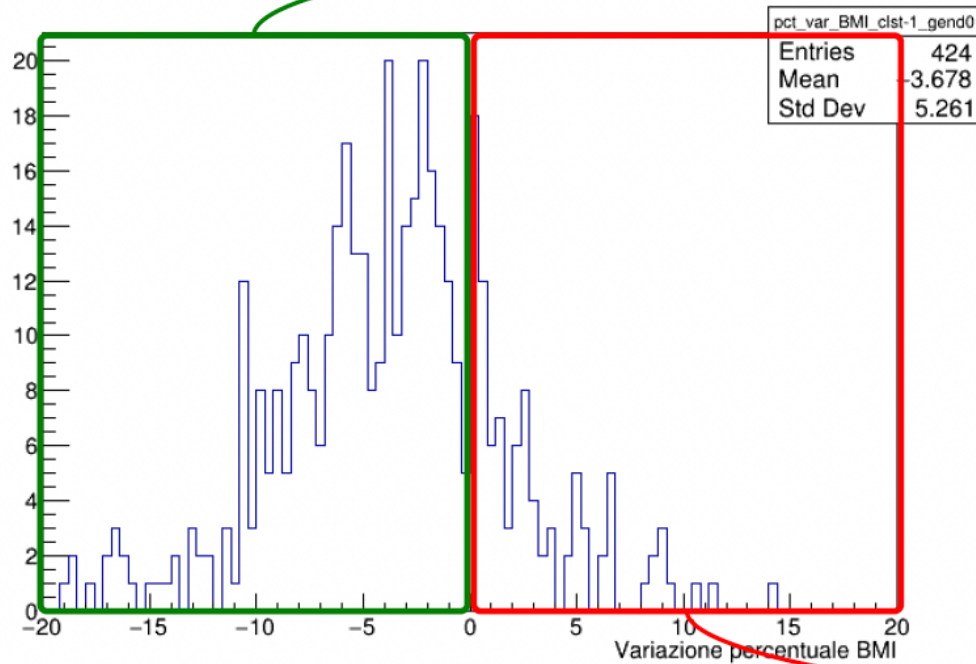
In media si arriva al risultato voluto in **circa 8 mesi**

Variazioni percentuali nel tempo

Consideriamo i valori di prima e ultima visita

Possiamo riempire un istogramma con la variazione percentuale definita come

$$\text{Variazione percentuale BMI} = \frac{\text{BMI finale} - \text{BMI iniziale}}{\text{BMI iniziale}}$$

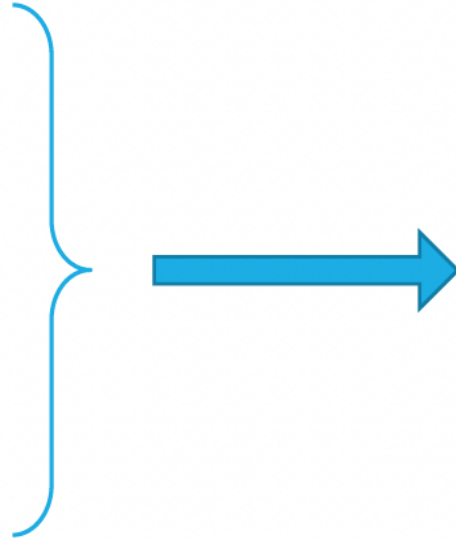


In base al cluster possiamo dire se la variazione rappresenti un miglioramento o un peggioramento della situazione del paziente

■ Variazione negativa
■ Variazione positiva

Conclusioni

- Presentate le variabili utilizzate
- Caratteristiche del campione
- Suddivisione in cluster
- Strategia di analisi



Nel prossimo talk verranno presentati i risultati in dettaglio

GRAZIE