Potential predictors of type-2 diabetes risk: machine learning, synthetic data and wearable health devices

Filippo Castiglione
National Research Council of Italy, Rome

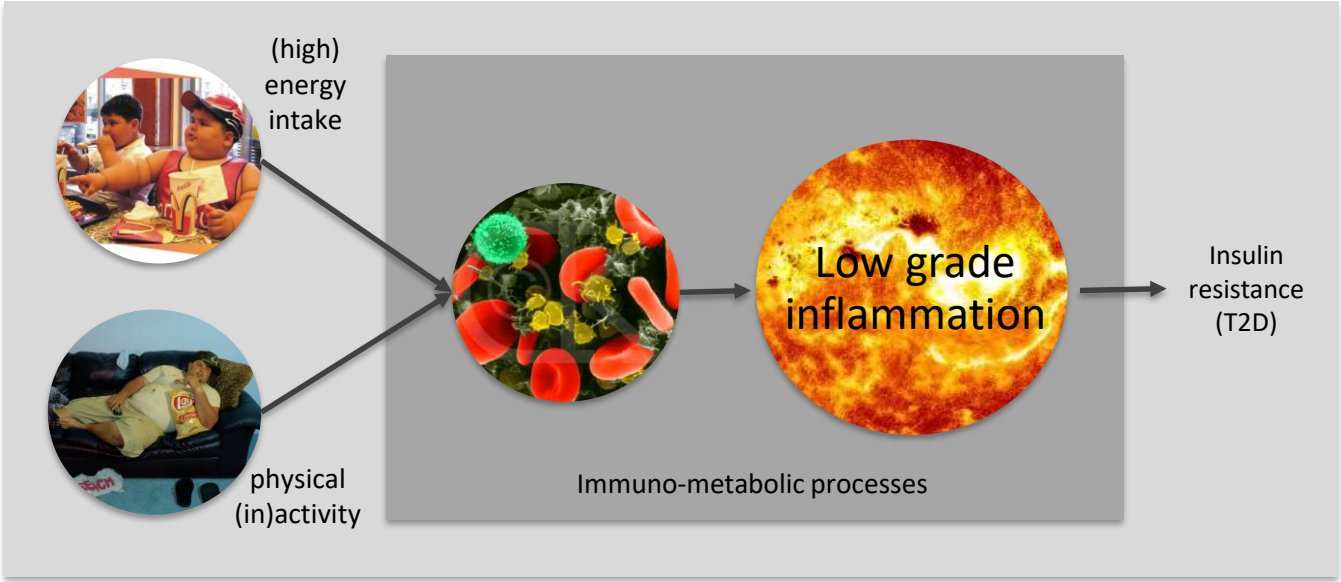
Workshop: the role of In Silico Medicine in diabetes
AIFA
Rome, May 26, 2022

T2D can be considered an auto-inflammatory disease

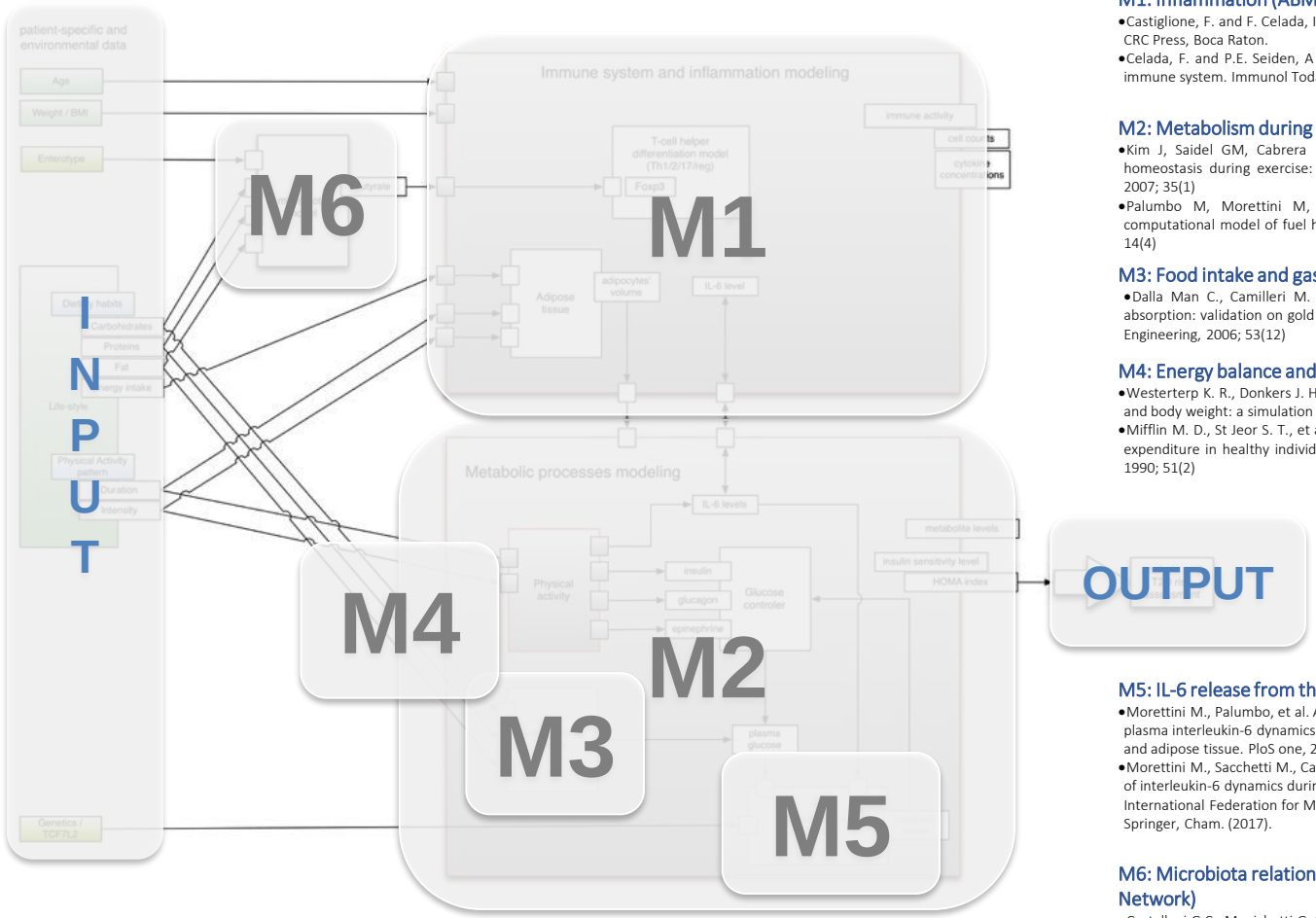
- evidence of inflammation markers in T2D (e.g., CRP, IL-6)
- clinical studies demonstrated the potential of use an anti-inflammatory drugs in T2D therapy

- Donath M., Shoelson S., Type 2 diabetes as an inflammatory disease, Nat. Rev. Immunol. 11: 98-107 (2011)
- Donath, M. Y., Boni-Schnetzler, M., Ellingsgaard, H. & Ehses, J. A. Islet inflammation impairs the pancreatic β-cell in type 2 diabetes. Physiology 24, 325–331 (2009)

Within the FP7-EU project MISSION-T2D we developed a **computational model** to study the emergence of T2D



Hotamisligil, G. S. Inflammation and metabolic disorders. Nature, 2006, 444, 860-867



M1: Inflammation (ABM)

- Castiglione, F. and F. Celada, Immune System Modeling and Simulation. 2015: CRC Press, Boca Raton.
- Celada, F. and P.E. Seiden, A computer model of cellular interactions in the immune system. Immunol Today, 1992; 13(2)

M2: Metabolism during exercise (ODEs)

- Kim J, Saidel GM, Cabrera ME. Multi-scale computational model of fuel homeostasis during exercise: effect of hormonal control. Ann Biomed Eng. 2007; 35(1)
- Palumbo M, Morettini M, et al. Personalizing physical exercise in a computational model of fuel homeostasis. PLoS computational biology. 2018; 14(4)

M3: Food intake and gastric emptying (ODEs)

- Dalla Man C., Camilleri M. & Cobelli C. A system model of oral glucose absorption: validation on gold standard data. IEEE Transactions on Biomedical Engineering, 2006; 53(12)

M4: Energy balance and fat gain/loss (ODEs)

- Westerterp K. R., Donkers J. H. & Fredrix E. W. Energy intake, physical activity and body weight: a simulation model. British Journal of Nutrition, 1995; 73(3)
- Mifflin M. D., St Jeor S. T., et al. A new predictive equation for resting energy expenditure in healthy individuals. The American journal of clinical nutrition, 1990; 51(2)

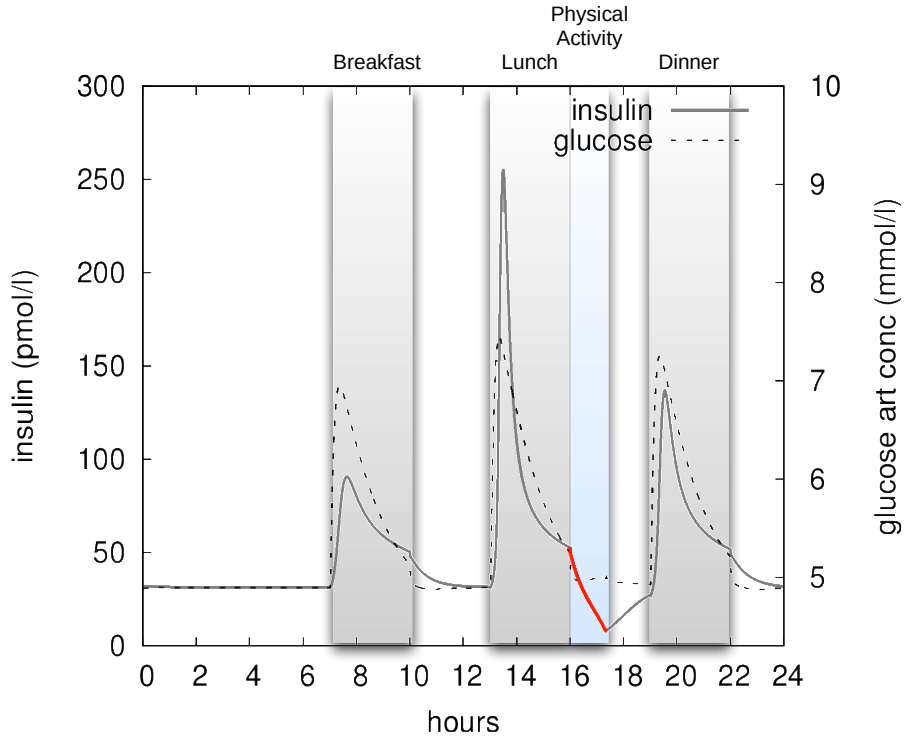
M5: IL-6 release from the skeletal muscle during PA (ODEs)

- Morettini M., Palumbo, et al. A system model of the effects of exercise on plasma interleukin-6 dynamics in healthy individuals: role of skeletal muscle and adipose tissue. PloS one, 2017; 12(7)
- Morettini M., Sacchetti M., Cappozzo, A., & Mazzà C. A mathematical model of interleukin-6 dynamics during exercise. In 6th European Conference of the International Federation for Medical and Biological Engineering (pp. 431-434). Springer, Cham. (2017).

M6: Microbiota relationship with inflammation (Multiplex Network)

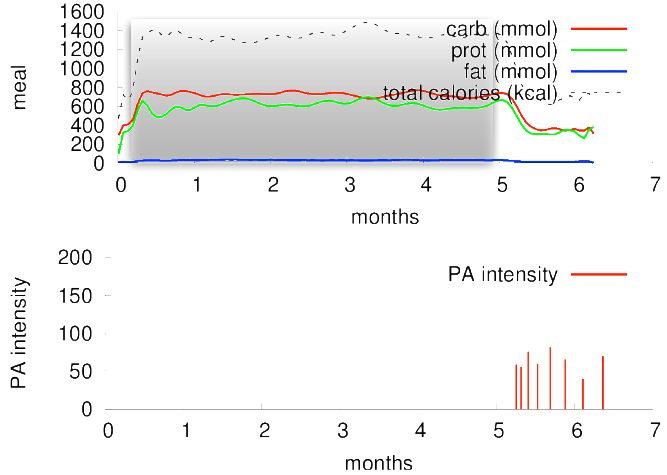
- Castellani G C., Menichetti G, Garagnani P, et al. Systems medicine of inflammaging, *Briefings in Bioinformatics*, 2016; 17(32016)

One day

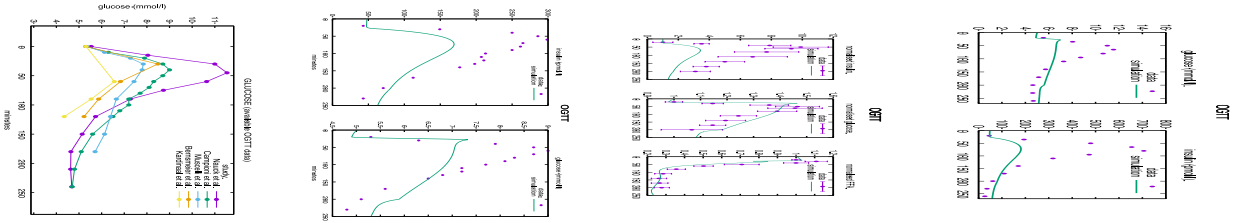

Physical Activity
Breakfast
Lunch
Dinner
insulin
glucose
insulin (pmol/l)
glucose art conc (mmol/l)
hours
0
50
100
150
200
250
300
5
6
7
8
9
10
0 2 4 6 8 10 12 14 16 18 20 22 24

6 months

Long period of high caloric intake followed by diet and PA



input



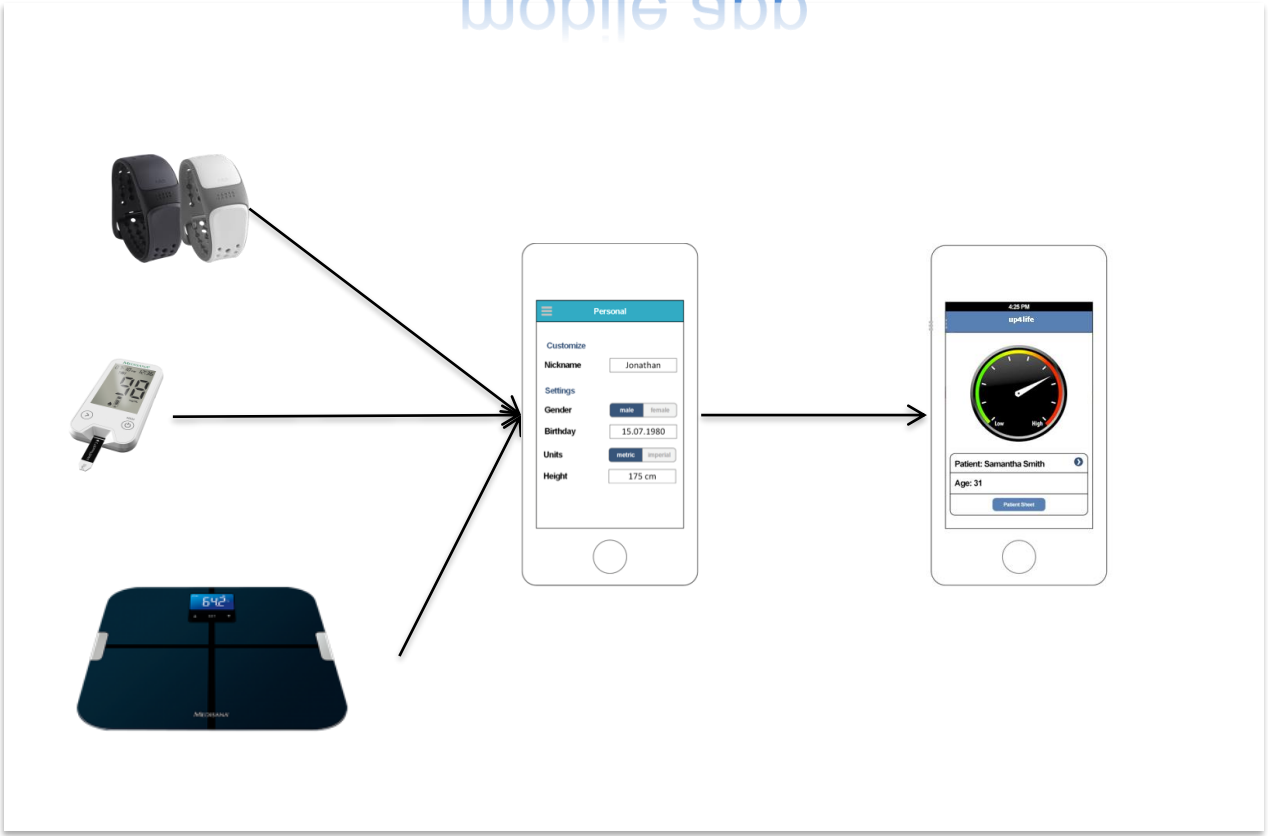
Validation

Physical activity

Lifestyle habits (diet + pa)

References	Input parameters	Output variables
Heilbronn [8]	age (37yr), weight (81.8kg), BMI ($27.5 kg \cdot m^{-2}$), CR diet (-25%kcal for 6 months), CREX diet (-12.5%kcal for 6 months, controllare pa), LCD diet (890kcal/day for 10 weeks)	weight, glucose, insulin
Kardinaal [7]	age (42.5yr), BMI ($24.1 kg \cdot m^{-2}$), weight (78.4kg), HFHC diet (+1300kcal/day)	weight, glucose, insulin
Wolfe [4]	weight (78kg), pa duration (60min), pa intensity (37%VO2max), VO2max ($40.5 mL^{-1} \cdot kg^{-1} \cdot min^{-1}$)	insulin, glucagon, epinephrine, FFA

mobile app


Personal
Customize
Nickname
Jonathan
Settings
Gender
male
female
Birthday
15.07.1980
Units
metric
imperial
Height
175 cm
up4life
Patient: Samantha Smith
Age: 31
Patient Sheet

MEDISANA

Download on the App Store

ANDROID APP ON Google Play

Data synthesis for ML

ANTHROPOMETRIC MEASURES

Sex	$\boldsymbol{S} \in \{\text{female, male}\}$
Age	$\boldsymbol{A} \in \{28,38,48,58,68\}$
Weight	$\boldsymbol{W} \in \{\text{underweight, normal, overweight}\}$
Height	$\boldsymbol{H} \in \{\text{short, average, tall}\}$

$$BMI_0 = W_0/H^2$$

PHYSICAL ACTIVITY

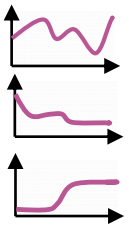
Number of sessions per week	$\boldsymbol{N_{PA}} \in \{0, 1, 2, 3\}$
Duration (mins)	$\boldsymbol{D_{PA}} \in \{\text{low} = 30, \text{medium} = 60, \text{high} = 90\}$
Intensity (% of VO_{2max})	$\boldsymbol{I_{PA}} \in \{\text{low} = 40, \text{high} = 60\}$

FOOD INTAKE (3 meals per day, breakfast, lunch, dinner)

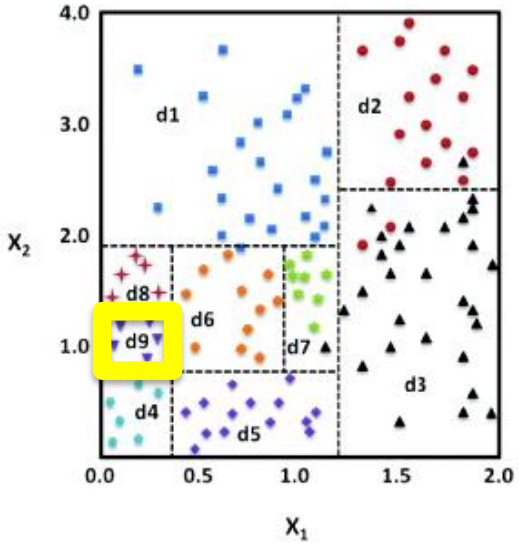
Carbohydrates (grams)	$\boldsymbol{C_{ME}} \in \{\text{low, med, high}\}$
Proteins (grams)	$\boldsymbol{P_{ME}} \in \{\text{low, med, high}\}$
Fats (grams)	$\boldsymbol{F_{ME}} \in \{\text{low, med, high}\}$

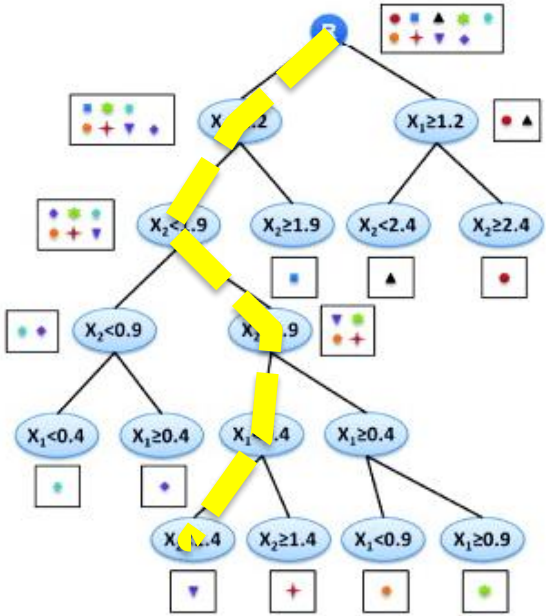
$$x = [S, A, BMI_0, (N_{PA}, D_{PA}, I_{PA}), (C_{ME}, P_{ME}, F_{ME})]$$
$$y = [BMI(t), GBL(t), TNF(t)]$$

$BMI(t)$ body mass index
$GBL(t)$ glucose base level (fasting glucose)
$TNF(t)$ inflammation level
t = six months



Decision tree

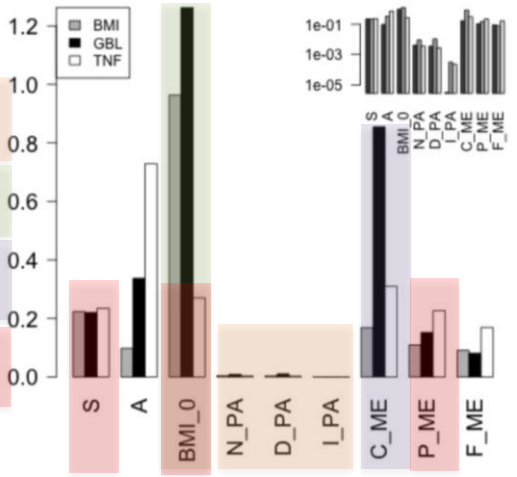




Variables' importance

Sensitivity analysis reveals that:

- the variables related to the physical activity (*i.e.*, N_{PA}, D_{PA} and I_{PA}) appear as the less important
- The most important variable for both the BMI and GBL is the initial value BMI_0
- GBL also strongly depends on the amount of carbohydrates in the diet (C_{ME})
- As for inflammation (TNF), the most important dependence is age (A) followed in order of importance by C_{ME} and then BMI_0



- H. Ishwaran, "Variable importance in binary regression trees and forests," Electron. J. Statist., vol. 1, pp. 519-537, Nov 2007.

Conclusions

- A tool which allows people's self-assessment based on lifestyle parameters remains the most powerful means to increase awareness of the risk of T2D

- The computational model can be used to feed a ML method which demonstrated to perform well to predict the risk of T2D using notably less computational resources, making it compliant for mobile devices

- The ML model can be used to compute optimal lifestyle patterns with respect to users' characteristics

Consiglio Nazionale delle Ricerche (IT)
Università di Bologna (IT)
Toegepast Natuurwetenschappelijk Onderzoek (NL)
University of Cambridge (UK)
Università degli Studi di Roma "Foro Italico" (IT)
University of Sheffield (UK)
Medisana Space Technologies GMbH (DE)

Thank you for your attention

Acknowledgments
M-C. Palumbo
P. Stolfi
P. Tieri