



## INTELLIGENZA ARTIFICIALE-MACHINE LEARNING

*ANALISI SU*

*IPERCOLESTEROLEMIA NEL PAZIENTE CON DIABETE TIPO 2: FENOTIPIZZAZIONE PREDITTIVA  
SUL RAGGIUNGIMENTO DEL TARGET COLESTEROLO LDL-C <100  
ENTRO 2 ANNI DALL'INIZIO DELLA TERAPIA IPOLIPEMIZZANTE*

 RuleX\*



\* Awarded "One of the Ten Most Disruptive Technologies" by MIT Sloan

**Gartner**\*

Recognized by Gartner as Cool Vendor in #AI Core Tech for 2020.

## TEAM DI PROGETTO

**RICCARDO CANDIDO**

*Vice Presidente AMD*

**GIACOMO GUAITA**

*Consigliere Nazionale AMD*

**NICOLETTA MUSACCHIO**

*Past President AMD*

**PAOLA PONZANI**

*Consigliere Nazionale AMD*

**PIERLUIGI SANTIN**

*Esperto Machine Learning Rulex®*

**RITA ZILICH**

*Mix-x*

## EXECUTIVE SUMMARY RISULTATI

- Questo studio conferma l'elevata percentuale di pazienti con LDL-C non a target entro 2 anni dall'inizio della terapia ipolipemizzante.
- Il rapporto fra LDL-C a inizio terapia e % di calo alla visita successiva differenzia (o caratterizza) i pazienti che raggiungono il target rispetto a quelli che non lo raggiungono. Il machine learning ha prodotto, per diversi range di LDL-C a inizio terapia ( $>100 \leq 125$ ,  $>125 \leq 150$ ,  $>150 \leq 175$ ,  $>175 \leq 200$ ,  $>200 \leq 250$ ), le riduzioni minime che devono essere ottenute entro la visita successiva per poter (probabilmente) raggiungere il target (LDL-C  $<100$ ) dopo 2 anni.
- **Lo studio conferma l'importanza dell'aggressività terapeutica.**
- **Lo studio suggerisce la necessità di puntare il prima possibile alla % di riduzione suggerita dal machine learning (vedi tabella) per garantire un buon risultato e la sua persistenza nel lungo periodo.**

Range di LDL-C a inizio terapia (T0)	Valori soglia indicati dai risultati del machine learning	
	Probabile target LDL-C dopo 2 anni: SI	Probabile target LDL-C dopo 2 anni: NO
<b><math>&gt;100 \leq 125</math></b>	Riduzione LDL-C alla visita successiva a T0 $>19\%$	Riduzione LDL-C alla visita successiva a T0 $<12\%$
<b><math>&gt;125 \leq 150</math></b>	Riduzione LDL-C alla visita successiva a T0 $>29\%$	Riduzione LDL-C alla visita successiva a T0 $<29\%$
<b><math>&gt;150 \leq 175</math></b>	Riduzione LDL-C alla visita successiva a T0 $>36\%$	Riduzione LDL-C alla visita successiva a T0 $<31\%$
<b><math>&gt;175 \leq 200</math></b>	Riduzione LDL-C alla visita successiva a T0 $>49\%$	Riduzione LDL-C alla visita successiva a T0 $<42\%$
<b><math>&gt;200 \leq 250</math></b>	Riduzione LDL-C alla visita successiva a T0 $>48\%$	Riduzione LDL-C alla visita successiva a T0 $<41\%$

## PREMESSA:

### L'IMPORTANZA DEL CONTROLLO DEL COLESTEROLO NEL PAZIENTE CON DIABETE TIPO 2

- L'ipercolesterolemia gioca un ruolo cruciale nello sviluppo della malattia aterosclerotica, particolarmente nei soggetti con diabete mellito.
- Più dell'80% dei pazienti con diabete tipo 2 (DM2) presentano un'alterazione del profilo lipidico.
- La maggior parte dei pazienti con DM2 presentano un rischio cardiovascolare (CV) elevato o molto elevato (dai dati degli Annali AMD: 78.5% sono a rischio molto elevato, 20.9% a rischio elevato e 0.6% a rischio moderato).
- Tra i pazienti con DM2 a rischio CV molto elevato, solo il 66% raggiunge livelli di LDL-C < 100 mg/dl e tra quelli a rischio elevato tale percentuale si riduce a poco più della metà (54.7%).
- Di converso, il 13% dei pazienti a rischio CV molto elevato ed il 9% di quelli a rischio elevato presentano valori di LDL-C  $\geq$  130 mg/dl.

## PREMESSA (2)

### L'IMPORTANZA DEL CONTROLLO DEL COLESTEROLO NEL PAZIENTE CON DIABETE TIPO 2

- Diversi studi (Costa J 2006; Kearney PM 2008) hanno dimostrato che l'intervento sul colesterolo LDL induce nei soggetti con diabete effetti almeno pari a quelli ottenuti nei non diabetici: riduzione della mortalità totale (9%), cardiovascolare e degli eventi cardiovascolari (21%).
- La riduzione sia del rischio relativo sia assoluto è significativa, indipendentemente dai livelli di LDL-C iniziale, ed è presente sia in prevenzione primaria che secondaria.
- Lo studio IMPROVE-IT (Cannon CP 2015 ) ha dimostrato che ezetimibe, quando aggiunto alla terapia con statina, comparato a placebo ha determinato un'ulteriore riduzione del livello di colesterolo LDL (54 mg/dl vs. 70 mg/dl) che si è accompagnata ad una ulteriore riduzione degli outcomes CV.
- A fronte di questi risultati, una bassa percentuale di pazienti raggiunge gli obiettivi di LDL-C indicati dalle linee guida ed elevato resta ancora il rischio CV residuo.

## OBIETTIVO CONOSCITIVO DELL'ANALISI

OBIETTIVO CONOSCITIVO PER QUESTA ANALISI DI MACHINE LEARNING:

**Identificare le differenti caratteristiche dei pazienti trattati per ipercolesterolemia che raggiungono il target lipidico (LDL-C<100\*) entro 2 anni dall'inizio della terapia ipolipemizzante rispetto a quelli che, nonostante il trattamento, non lo raggiungono, attraverso la creazione di un modello conoscitivo/predittivo.**

**Questo allo scopo di orientare in modo preventivo l'atteggiamento terapeutico del medico.**

*\*Il target LDL-C è stato fissato a 100 in quanto le linee guida che ne hanno variato le soglie sono state pubblicate a metà del 2017 e il database ha raccolto dati a partire dal 2005, periodo in cui il target LDL-C per il paziente diabetico ad alto rischio era fissato a 100. Inoltre, per poter disporre di un'osservazione di almeno 2 anni dopo l'inizio della terapia, visto che il database finisce a metà 2019, sono stati selezionati i pazienti che hanno cominciato la terapia entro la metà del 2017, periodo in cui erano ancora vigenti le linee guida precedenti.*

## I DATI UTILIZZATI

Per questa analisi è stato utilizzato il

### **Database Annali AMD 2019**

(anni dal 01/01/2005 al 30/06/2019)

***Il database contiene:***

***1.437.260 pazienti, di cui 1.207.597 pazienti con almeno una visita e***

***un totale di 11.254.635 visite.***

LA NUMEROSITÀ DEL CAMPIONE RISULTANTE DOPO LA PULIZIA DEI DATI È PARI A

**1.194.005 INDIVIDUI**

È possibile visualizzare gli indicatori raccolti nel database di AMD al seguente link: <https://aemmedi.it/annali-amd/> (alle voci: STRUMENTI, lista indicatori AMD)

## APPROCCIO ALL'ANALISI

LA CREAZIONE DEL MODELLO PREDITTIVO & CONOSCITIVO HA IMPLICATO LO SVOLGIMENTO DELLE SEGUENTI ATTIVITÀ, CHE HANNO INTEGRATO FASI DI BUSINESS INTELLIGENCE (FOTOGRAFIE DELLA SITUAZIONE) E FASI DI MACHINE LEARNING (ALGORITMI PREDITTIVI):

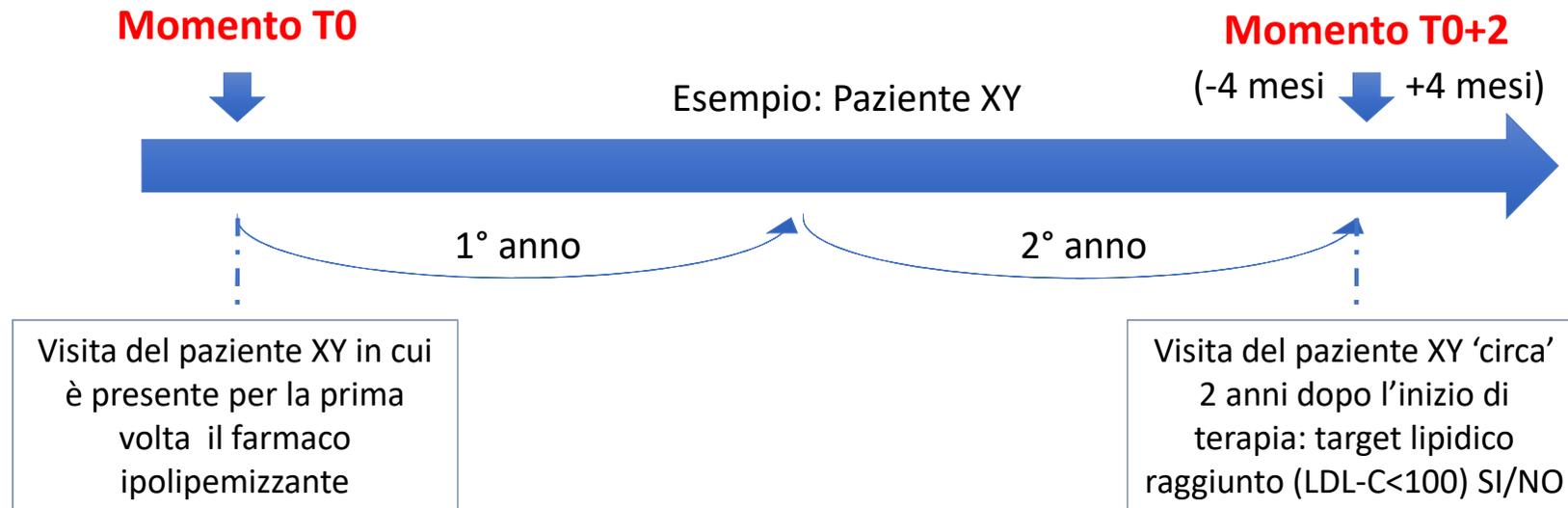
### 1. FOTOGRAFIA DELLA SITUAZIONE:

- a) Caratteristiche/numerosità della coorte (età, range di LDL, grado di rischio cardiovascolare);
- b) Fotografia dei pazienti a inizio terapia ipolipemizzante e dopo 2 anni (da varie prospettive: con esplicitazione dei livelli di LDL-C, del grado di rischio cardiovascolare, dei livelli di glicata);
- c) Fotografia sull'andamento nell'utilizzo dei farmaci ipolipemizzanti negli anni: per classi di farmaci (statine, fibrati, omega3, ezetimibe, altro) e per tipologia di statina (atorvastatina, simvastatina, atorvastatina, pravastatina, fluvastatina, lovastatina, altro);
- d) Fotografia con la distribuzione dei pazienti negli anni, in base a range di LDL-C e di grado di rischio cardiovascolare (per la coorte a 'inizio terapia ipolipemizzante' e per la stessa coorte al momento '2 anni dopo l'inizio della terapia ipolipemizzante').

### 2. ANALISI DI MACHINE LEARNING:

- a) Creazione dei modelli predittivi (prima fase): esplicitazione delle variabili più rilevanti per la fenotipizzazione dei pazienti (in relazione all'obiettivo conoscitivo, ovvero il raggiungimento o meno del target di LDL-C <100 entro 2 anni da inizio terapia);
- b) Tracciamento dell'andamento nel tempo dell'LDL-C nei pazienti trattati, prima e dopo l'inizio della terapia ipolipemizzante;
- c) Creazione dei modelli predittivi (seconda fase);
- d) VERIFICA DEI RISULTATI DEL MACHINE LEARNING: EVIDENZE SULLE DINAMICHE DI ABBASSAMENTO DELL'LDL-C CORRELATE ALLA POSSIBILITÀ DI RAGGIUNGERE IL TARGET 2 ANNI DOPO L'INIZIO DELLA TERAPIA IPOLIPEMIZZANTE.

## CRITERI D'INCLUSIONE



Sono stati individuati i pazienti per cui, ad un certo punto della loro storia, è presente una prescrizione di farmaco ipolipemizzante (definito 'momento T0') ed è contestualmente presente la misurazione del valore dell'LDL-C. Si verifica poi se dopo 2 anni (circa) quegli stessi pazienti sono a target lipidico [per quelli per cui si ha una visita nell'intorno dei 2 anni successivi ('momento T0+2') & si ha la misurazione di LDL-C] oppure no.

FOTOGRAFIA DELLA SITUAZIONE  
(RELATIVA AI PAZIENTI DELLA COORTE)

## NUMEROSITÀ PAZIENTI DELLA COORTE PER FASCE DI ETÀ E SESSO

Fascia Età a T0	F	M	Totale
sotto i 40	0.11% 181	0.27% 453	0.38% 634
da 40 a 49	1.64% 2,767	3.27% 5,523	4.90% 8,290
da 50 a 59	7.74% 13,084	11.68% 19,735	19.42% 32,819
da 60 a 69	16.20% 27,376	21.29% 35,986	37.49% 63,362
da 70 a 79	15.05% 25,430	15.68% 26,498	30.72% 51,928
80 o più	4.03% 6,806	3.06% 5,174	7.09% 11,980
Totale complessivo	44.76% 75,644	55.24% 93,369	100.00% 169,013

## NUMEROSITÀ PAZIENTI PER FASCE DI ETÀ/RANGE DI LDL-C (A T0 E A T0+2)

T0

T0+2

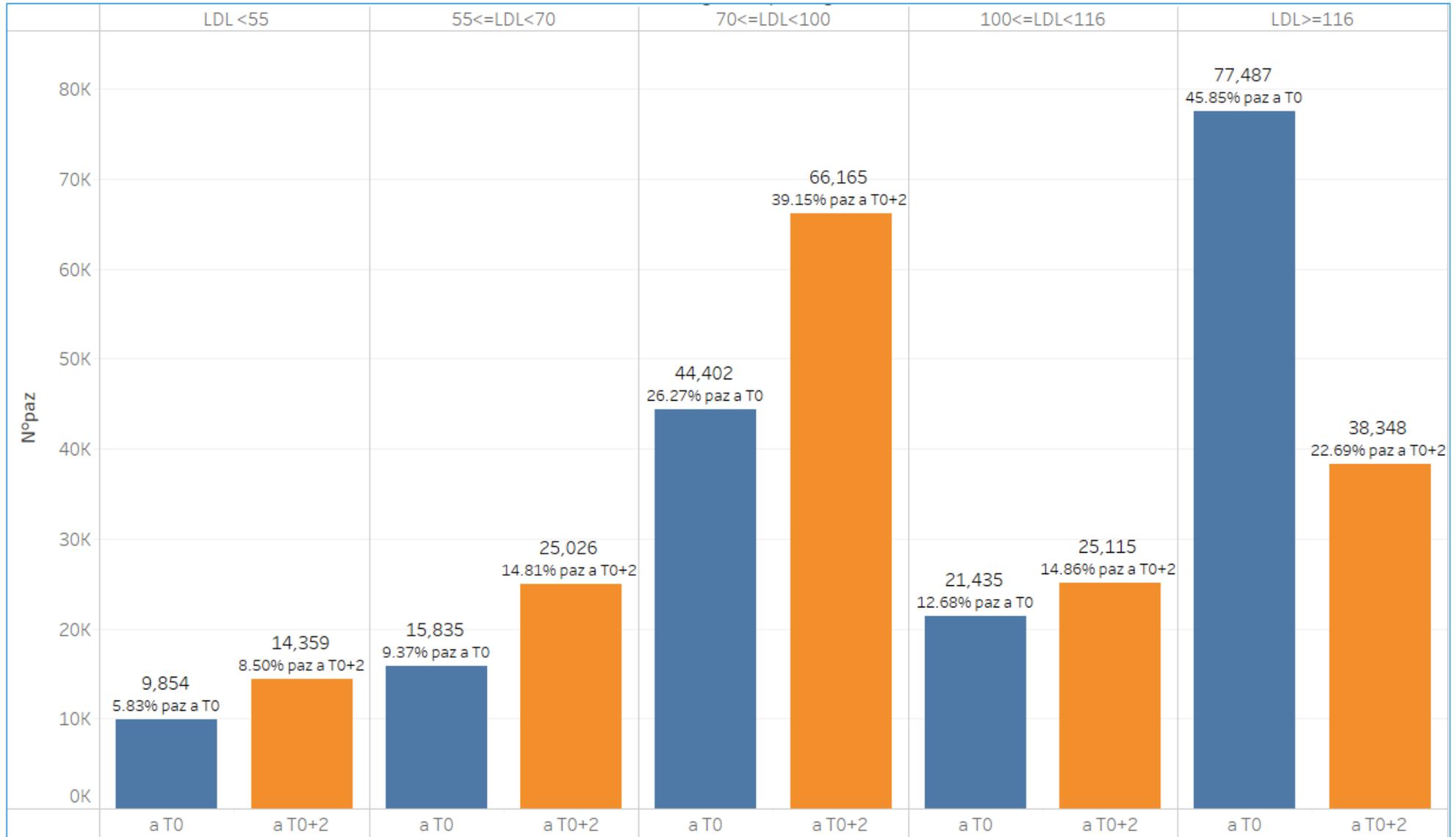
Fascia Età a T0	T0					Totale	T0+2					Totale
	LDL <55	55<=LDL<70	70<=LDL<100	100<=LDL<116	LDL >=116		LDL <55	55<=LDL<70	70<=LDL<100	100<=LDL<116	LDL >=116	
sotto i 40	31 0.02%	34 0.02%	87 0.05%	62 0.04%	420 0.25%	634 0.38%	47 0.03%	65 0.04%	215 0.13%	90 0.05%	217 0.13%	634 0.38%
da 40 a 49	350 0.21%	518 0.31%	1,414 0.84%	809 0.48%	5,199 3.08%	8,290 4.90%	547 0.32%	988 0.58%	2,968 1.76%	1,254 0.74%	2,533 1.50%	8,290 4.90%
da 50 a 59	1,500 0.89%	2,405 1.42%	7,053 4.17%	3,785 2.24%	18,076 10.70%	32,819 19.42%	2,305 1.36%	4,356 2.58%	12,671 7.50%	5,127 3.03%	8,360 4.95%	32,819 19.42%
da 60 a 69	3,540 2.09%	5,880 3.48%	16,706 9.88%	8,090 4.79%	29,146 17.24%	63,362 37.49%	5,419 3.21%	9,615 5.69%	25,001 14.79%	9,445 5.59%	13,882 8.21%	63,362 37.49%
da 70 a 79	3,431 2.03%	5,466 3.23%	15,262 9.03%	7,107 4.21%	20,662 12.23%	51,928 30.72%	4,742 2.81%	8,069 4.77%	20,714 12.26%	7,556 4.47%	10,847 6.42%	51,928 30.72%
80 o più	1,002 0.59%	1,532 0.91%	3,880 2.30%	1,582 0.94%	3,984 2.36%	11,980 7.09%	1,299 0.77%	1,933 1.14%	4,596 2.72%	1,643 0.97%	2,509 1.48%	11,980 7.09%
Totale complessivo	9,854 5.83%	15,835 9.37%	44,402 26.27%	21,435 12.68%	77,487 45.85%	169,013 100.00%	14,359 8.50%	25,026 14.81%	66,165 39.15%	25,115 14.86%	38,348 22.69%	169,013 100.00%

## NUMEROSITÀ PAZIENTI PER RANGE DI LDL E CV RISK\* (A T0 E A T0+2)

<b>T0</b>							<b>T0+2</b>					
CV RISK	LDL <55	55<=LDL<70	70<=LDL<100	100<=LDL<116	LDL>=116	Totale	LDL <55	55<=LDL<70	70<=LDL<100	100<=LDL<116	LDL>=116	Totale
<b>Alto*</b>	205 3.02%	377 5.56%	1,342 19.78%	805 11.87%	4,055 59.77%	6,784 100.00%	73 5.61%	157 12.06%	467 35.87%	197 15.13%	408 31.34%	1,302 100.00%
<b>Molto alto*</b>	9,649 5.95%	15,458 9.53%	43,060 26.54%	20,630 12.72%	73,432 45.26%	162,229 100.00%	14,286 8.52%	24,869 14.83%	65,698 39.17%	24,918 14.86%	37,940 22.62%	167,711 100.00%
<b>Totale complessivo</b>	9,854 5.83%	15,835 9.37%	44,402 26.27%	21,435 12.68%	77,487 45.85%	169,013 100.00%	14,359 8.50%	25,026 14.81%	66,165 39.15%	25,115 14.86%	38,348 22.69%	169,013 100.00%

\*GIC Volume 22 | Suppl. 1 al n. 4 | Aprile 2021

## NUMEROSITÀ PAZIENTI PER RANGE\* DI LDL-C (A T0 E A T0+2)



\*range riferiti alle nuove linee guida per il trattamento del colesterolo nel paziente con diabete

## NUMEROSITÀ PAZIENTI PER RANGE DI GLICATA E LDL-C\* (A T0 E A T0+2)

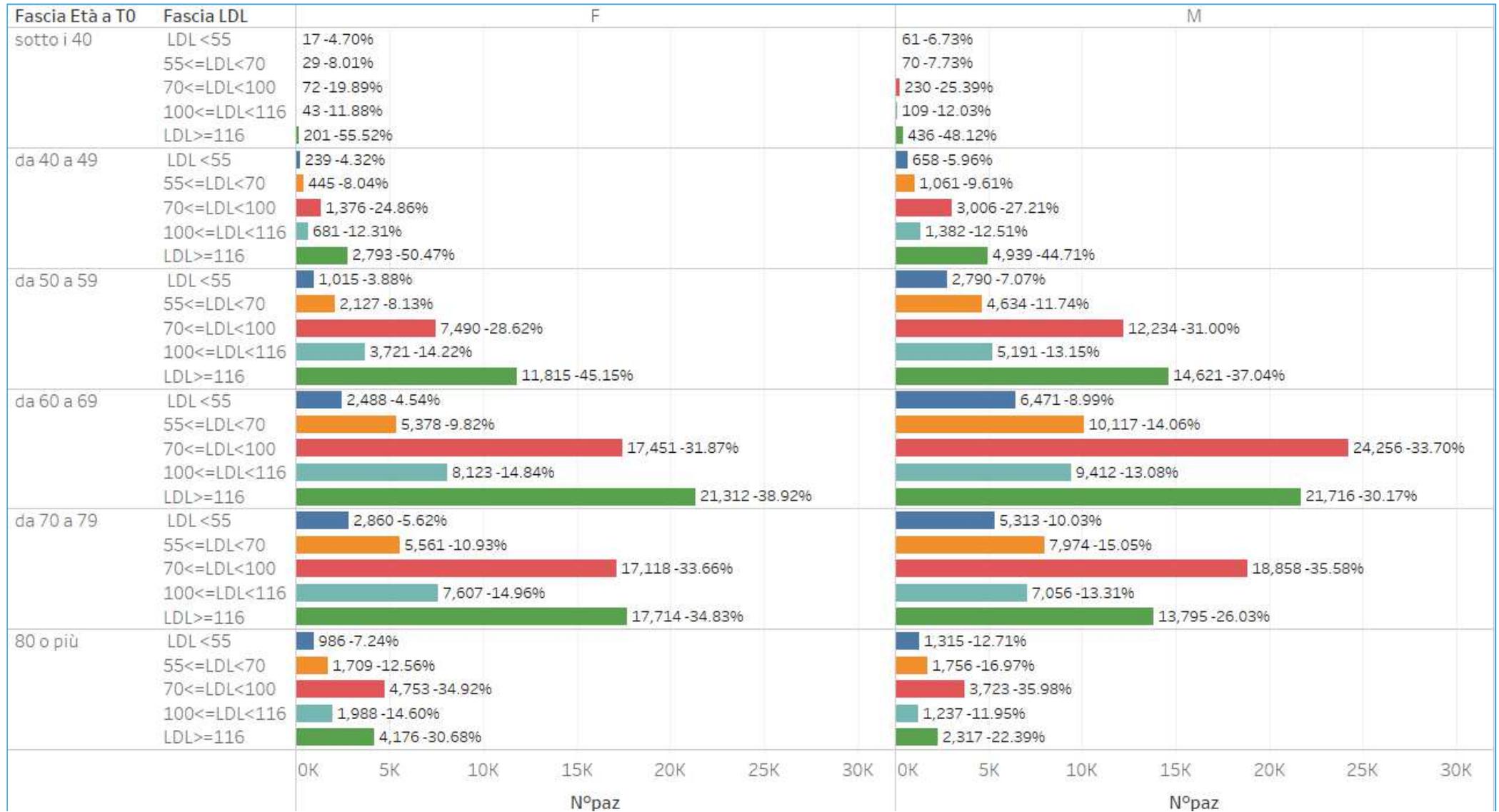
T0

T0+2

Fascia Glicata a T0	LDL <55	55<=LDL<70	70<=LDL<100	100<=LDL<116	LDL>=116	Totale	LDL <55	55<=LDL<70	70<=LDL<100	100<=LDL<116	LDL>=116	Totale
Glicata<7	3,772 2.23%	6,590 3.90%	18,550 10.98%	9,145 5.41%	33,552 19.85%	71,609 42.37%	5,502 3.26%	10,163 6.01%	28,392 16.80%	10,997 6.51%	16,555 9.80%	71,609 42.37%
7<=Glicata<8	2,903 1.72%	4,693 2.78%	13,372 7.91%	6,403 3.79%	21,516 12.73%	48,887 28.92%	4,318 2.55%	7,527 4.45%	19,231 11.38%	7,224 4.27%	10,587 6.26%	48,887 28.92%
8<=Glicata<9	1,560 0.92%	2,362 1.40%	6,806 4.03%	3,127 1.85%	10,638 6.29%	24,493 14.49%	2,262 1.34%	3,759 2.22%	9,490 5.61%	3,526 2.09%	5,456 3.23%	24,493 14.49%
Glicata>=9	1,619 0.96%	2,190 1.30%	5,674 3.36%	2,760 1.63%	11,781 6.97%	24,024 14.21%	2,277 1.35%	3,577 2.12%	9,052 5.36%	3,368 1.99%	5,750 3.40%	24,024 14.21%
Totale complessivo	9,854 5.83%	15,835 9.37%	44,402 26.27%	21,435 12.68%	77,487 45.85%	169,013 100.00%	14,359 8.50%	25,026 14.81%	66,165 39.15%	25,115 14.86%	38,348 22.69%	169,013 100.00%

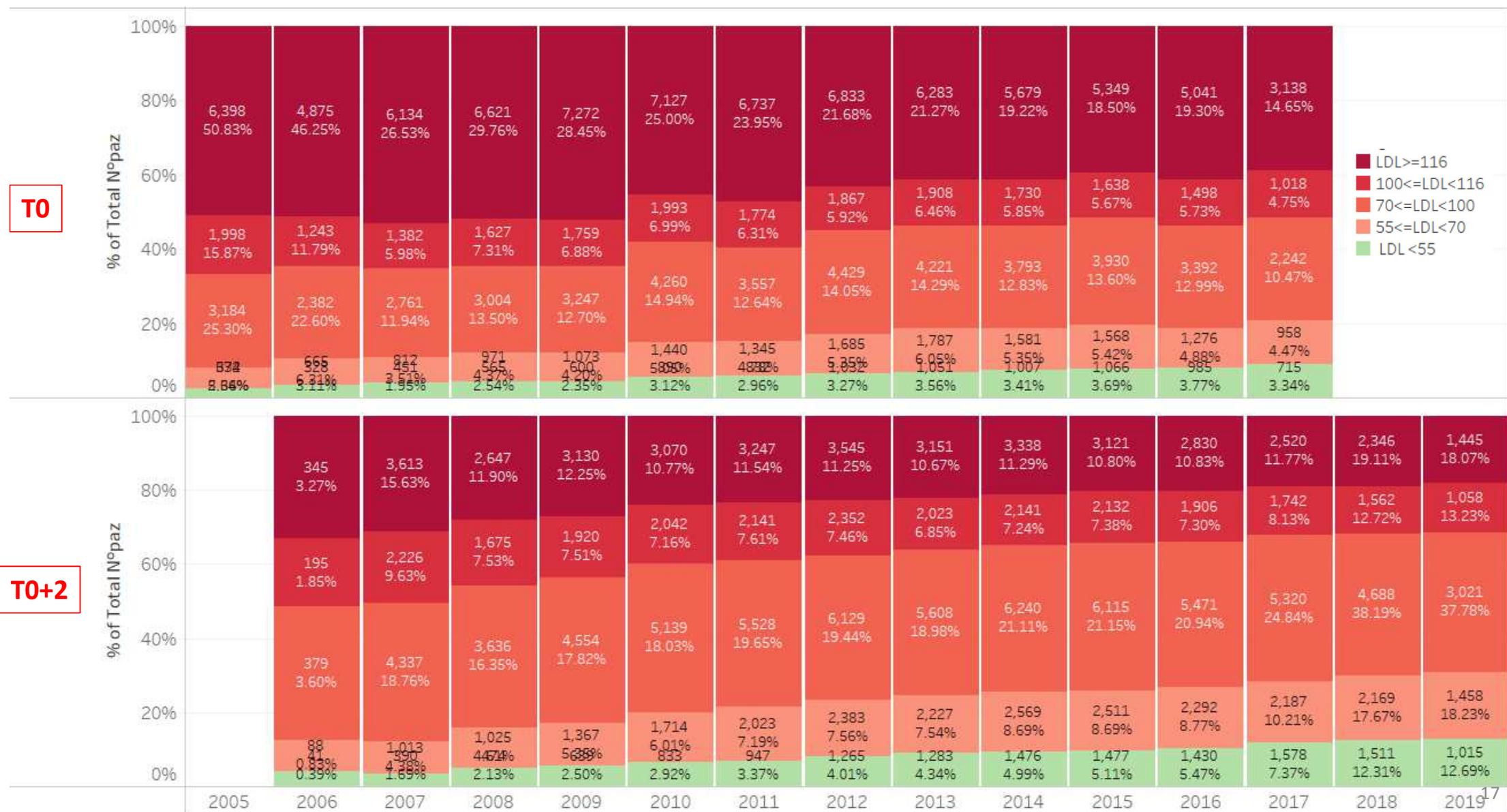
*\*gli incroci corrispondenti ai riquadri più scuri hanno una maggiore numerosità di individui*

## NUMEROSITÀ PAZIENTI\* PER RANGE DI LDL-C, ETÀ E SESSO



\*I numeri riportati si riferiscono a: n. pazienti - % pazienti rispetto al totale pazienti di quella fascia di età

## DISTRIBUZIONE PAZIENTI PER ANNO, PER RANGE DI LDL-C E CV RISK (A T0 E A T0+2)



FOTOGRAFIA DELLA SITUAZIONE  
UTILIZZO DEI FARMACI IPOLIPEMIZZANTI NEGLI ANNI

# NUMEROSITÀ UTILIZZO DELLE DIVERSE CLASSI DI FARMACI IPOLIPEMIZZANTI\*, PER ANNO

TIPO FARMACO	Con Statina	Data															Totale co..
		2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	
EZETIMIBE	no	0.00% 0	0.00% 2	0.03% 43	0.08% 164	0.09% 222	0.10% 276	0.23% 718	0.70% 2,388	1.09% 4,014	1.35% 5,207	1.61% 6,362	1.86% 7,506	2.17% 9,054	2.40% 10,809	2.59% 6,018	1.21% 52,783
	si	0.09% 70	1.28% 1,450	3.28% 5,084	4.23% 8,332	4.79% 11,523	5.13% 14,480	5.49% 16,914	5.90% 20,034	6.37% 23,533	6.84% 26,318	7.49% 29,681	8.05% 32,455	8.97% 37,372	9.88% 44,521	11.21% 26,051	6.82% 297,818
FIBRATI	no	4.96% 3,773	4.51% 5,091	4.36% 6,772	4.34% 8,559	4.12% 9,920	3.80% 10,727	3.52% 10,857	3.37% 11,430	3.31% 12,246	3.37% 12,940	3.29% 13,037	3.29% 13,260	3.23% 13,447	3.28% 14,764	3.27% 7,602	3.54% 154,425
	si	0.15% 112	0.15% 174	0.23% 353	0.31% 604	0.37% 883	0.47% 1,337	0.58% 1,802	0.67% 2,286	0.92% 3,408	1.11% 4,266	1.30% 5,163	1.51% 6,091	1.73% 7,202	1.92% 8,671	2.02% 4,701	1.08% 47,053
OMEGA-3	no	4.24% 3,223	3.53% 3,980	3.31% 5,131	3.04% 5,992	2.89% 6,965	2.78% 7,852	2.73% 8,402	2.63% 8,940	2.47% 9,120	2.25% 8,652	2.15% 8,510	2.17% 8,767	2.34% 9,733	2.49% 11,220	2.48% 5,758	2.57% 112,245
	si	8.21% 6,241	8.36% 9,436	8.57% 13,303	8.93% 17,613	9.20% 22,144	9.65% 27,265	10.05% 30,976	10.26% 34,841	9.60% 35,495	8.68% 33,375	8.31% 32,928	8.17% 32,949	8.40% 34,973	8.61% 38,773	8.64% 20,091	8.94% 390,403
STATINE	no	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0
	si	90.33% 68,695	91.69% 103,518	92.13% 142,973	92.47% 182,294	92.95% 223,731	93.42% 263,840	93.67% 288,576	93.60% 317,711	93.51% 345,712	93.48% 359,447	93.46% 370,338	93.24% 375,887	92.93% 387,085	92.58% 417,050	92.46% 214,970	93.05% 4,061,827
ALTRO	no	0.84% 639	0.64% 719	0.57% 877	0.47% 934	0.38% 920	0.32% 911	0.28% 862	0.24% 798	0.21% 784	0.17% 645	0.15% 608	0.15% 602	0.13% 541	0.13% 586	0.12% 290	0.25% 10,716
	si	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0.00% 0

\*Sono state calcolate le visite per anno in cui è presente una certa classe di farmaci (indipendentemente dal fatto che i farmaci siano presenti da soli o in associazione con altri).

## NUMEROSITÀ UTILIZZO DELLE DIVERSE TIPOLOGIE DI STATINE, PER ANNO

Tipo di statina	Data															Totale c..
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	
Atorvastatina	41.12% 28,245	39.24% 40,621	36.90% 52,762	35.64% 64,975	35.08% 78,493	33.95% 89,566	33.93% 97,904	35.28% 112,101	38.07% 131,600	40.17% 144,397	41.92% 155,252	43.89% 164,960	46.27% 179,086	47.74% 199,097	47.86% 102,884	40.42% 1,641,943
Simvastatina	29.81% 20,475	28.48% 29,485	30.13% 43,081	32.15% 58,612	33.05% 73,954	33.75% 89,043	34.88% 100,648	35.66% 113,290	35.09% 121,301	34.47% 123,918	33.87% 125,421	33.10% 124,401	31.83% 123,205	30.81% 128,513	29.99% 64,473	32.99% 1,339,820
Rosuvastatina	9.91% 6,808	15.03% 15,556	17.69% 25,290	19.14% 34,888	20.92% 46,804	22.62% 59,677	22.52% 64,991	21.17% 67,248	19.44% 67,190	18.22% 65,479	17.27% 63,949	16.13% 60,645	15.07% 58,345	14.35% 59,836	14.08% 30,267	17.90% 726,973
Pravastatina	11.89% 8,167	9.70% 10,039	7.99% 11,425	6.71% 12,227	5.64% 12,619	4.87% 12,860	4.41% 12,720	4.05% 12,853	3.78% 13,072	3.55% 12,775	3.40% 12,577	3.29% 12,350	3.21% 12,437	3.06% 12,780	2.97% 6,384	4.32% 175,285
Fluvastatina	6.65% 4,568	6.16% 6,381	5.47% 7,826	4.40% 8,027	3.40% 7,614	2.76% 7,274	2.23% 6,441	1.78% 5,660	1.48% 5,105	1.27% 4,566	1.11% 4,103	1.00% 3,770	0.85% 3,299	0.73% 3,028	0.66% 1,409	1.95% 79,071
Lovastatina	0.12% 80	0.90% 936	1.41% 2,011	1.55% 2,826	1.48% 3,314	1.60% 4,209	1.55% 4,476	1.54% 4,908	1.59% 5,485	1.66% 5,962	1.75% 6,470	1.81% 6,800	1.86% 7,207	1.93% 8,066	1.91% 4,116	1.65% 66,866
Altro (mix)	0.51% 352	0.48% 500	0.40% 578	0.41% 739	0.42% 933	0.46% 1,211	0.48% 1,396	0.52% 1,651	0.57% 1,959	0.65% 2,350	0.69% 2,566	0.79% 2,961	0.91% 3,506	1.37% 5,730	2.53% 5,437	0.78% 31,869
<b>Totale complessivo</b>	<b>100.00%</b> <b>68,695</b>	<b>100.00%</b> <b>103,518</b>	<b>100.00%</b> <b>142,973</b>	<b>100.00%</b> <b>182,294</b>	<b>100.00%</b> <b>223,731</b>	<b>100.00%</b> <b>263,840</b>	<b>100.00%</b> <b>288,576</b>	<b>100.00%</b> <b>317,711</b>	<b>100.00%</b> <b>345,712</b>	<b>100.00%</b> <b>359,447</b>	<b>100.00%</b> <b>370,338</b>	<b>100.00%</b> <b>375,887</b>	<b>100.00%</b> <b>387,085</b>	<b>100.00%</b> <b>417,050</b>	<b>100.00%</b> <b>214,970</b>	<b>100.00%</b> <b>4,061,827</b>

## ANALISI MACHINE LEARNING: PRIMA FASE

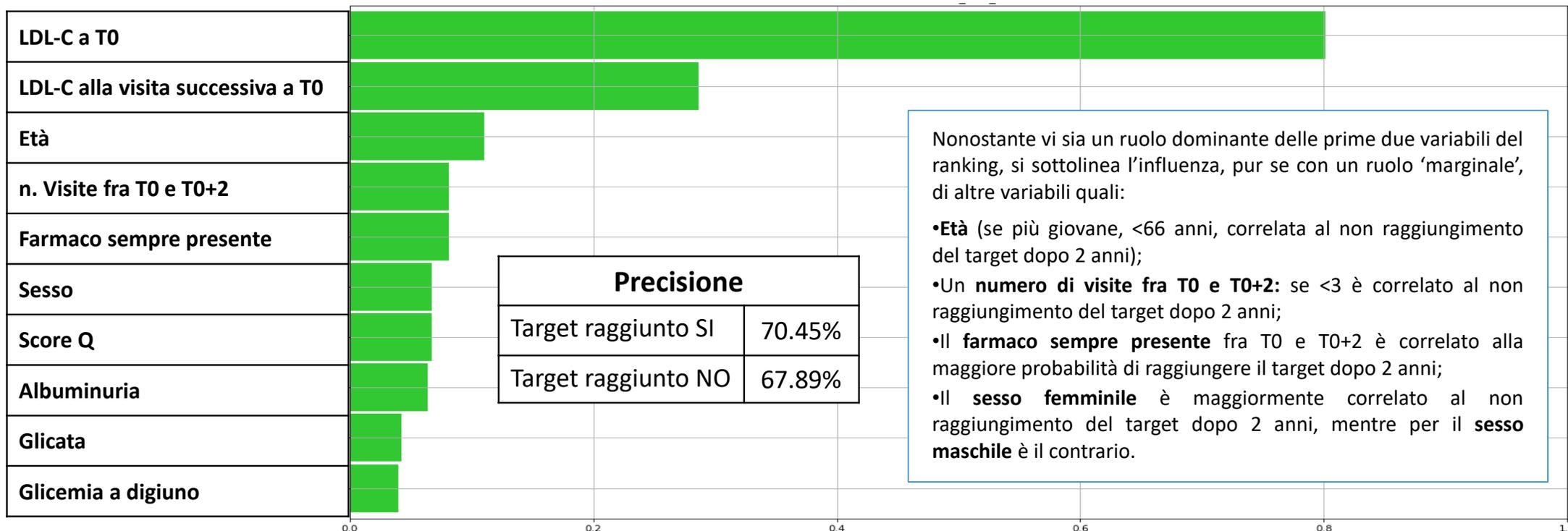
MODELLO PREDITTIVO/CONOSCITIVO (FENOTIPIZZAZIONE)

PROBABILITÀ DI RAGGIUNGERE IL TARGET DI LDL-C <100 ENTRO 2 ANNI DA INIZIO TERAPIA

# PROBABILE RAGGIUNGIMENTO DEL TARGET DI LDL-C ENTRO 2 ANNI DA INIZIO TERAPIA?

## IL RUOLO CHIAVE DEL RAPPORTO FRA LDL-C A INIZIO TERAPIA E LA SUA VARIAZIONE ALLA VISITA SUCCESSIVA

Fra le variabili maggiormente correlate alla possibilità che il paziente sia a target dopo 2 anni da inizio terapia ipolipemizzante, il modello di machine learning ha evidenziato la forte dominanza del rapporto fra LDL-C a inizio terapia e il suo valore alla visita successiva. Il machine learning ha suggerito che vi sono dei 'valori soglia' di riduzione dell'LDL-C entro la visita successiva all'inizio della terapia; tali valori soglia sono correlati alla maggiore/minore probabilità di raggiungimento del target dopo 2 anni da inizio terapia. Su questi valori, in quanto parametri chiave, sono stati fatti degli approfondimenti nell'analisi presentati alle slide successive.



Nonostante vi sia un ruolo dominante delle prime due variabili del ranking, si sottolinea l'influenza, pur se con un ruolo 'marginale', di altre variabili quali:

- **Età** (se più giovane, <66 anni, correlata al non raggiungimento del target dopo 2 anni);
- Un **numero di visite fra T0 e T0+2**: se <3 è correlato al non raggiungimento del target dopo 2 anni;
- Il **farmaco sempre presente** fra T0 e T0+2 è correlato alla maggiore probabilità di raggiungere il target dopo 2 anni;
- Il **sexo femminile** è maggiormente correlato al non raggiungimento del target dopo 2 anni, mentre per il **sexo maschile** è il contrario.

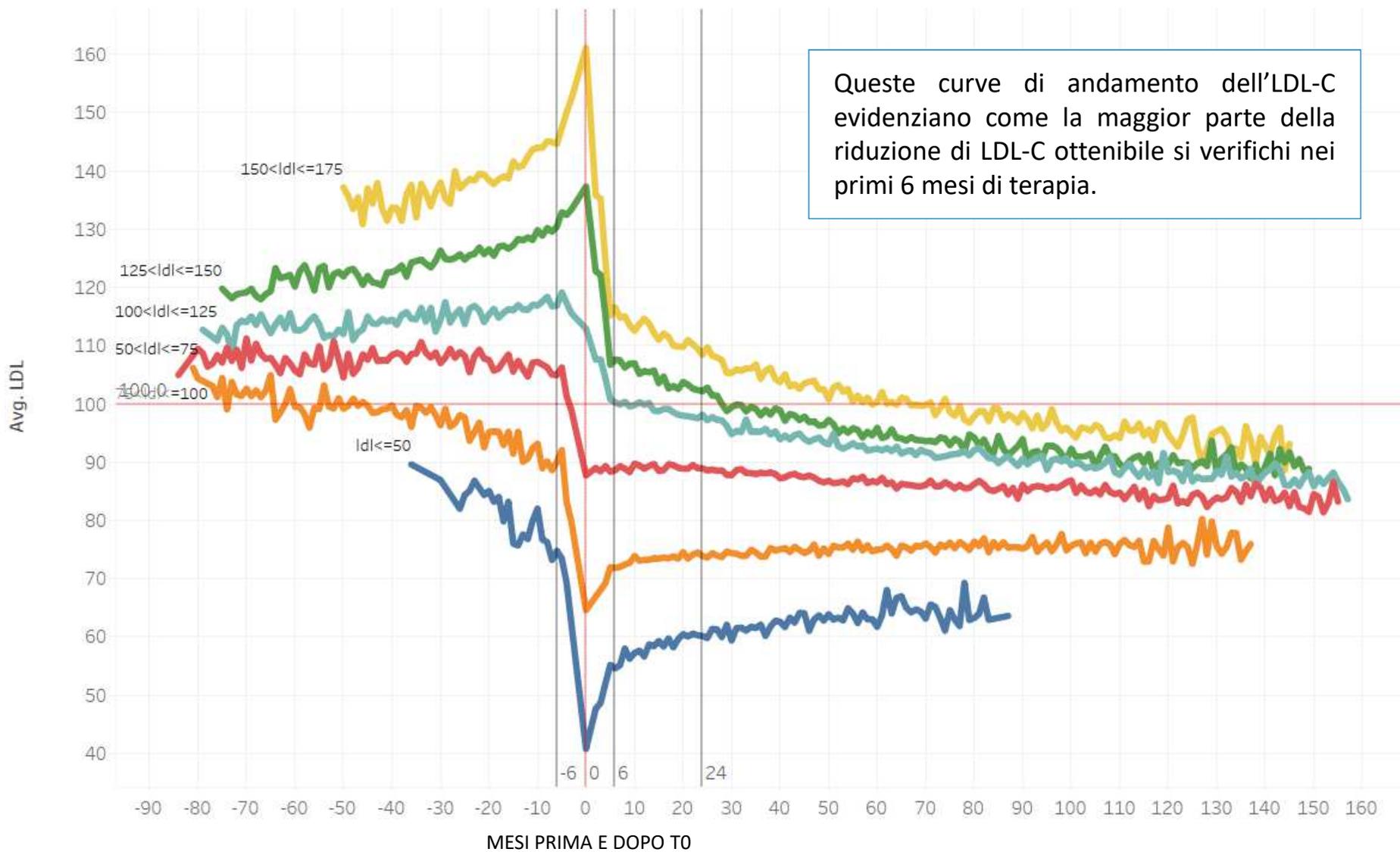
La lunghezza delle barrette verdi si riferisce all'importanza relativa fra le variabili (Thorsen-Meyer; Hans-Christian; Nielsen, Annelaura B. et al / Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records. In: The Lancet Digital Health. 2020 ; Vol. 2, No. 4. pp. e179–91.)

## APPROFONDIMENTO DELL'ANALISI

- IL RUOLO DETERMINANTE EVIDENZIATO DAL MACHINE LEARNING DEL RAPPORTO FRA LDL-C A INIZIO TERAPIA E IL SUO VALORE ALLA VISITA SUCCESSIVA (RISPETTO ALLA PROBABILITÀ DI RAGGIUNGERE IL TARGET DOPO 2 ANNI) HA GUIDATO UN APPROFONDIMENTO DELL'ANALISI;
- SONO STATE DISEGNATE LE CURVE CHE RIPORTANO L'ANDAMENTO MEDIO DELL'LDL-C NEI PAZIENTI DELLA COORTE, PRIMA E DOPO T0, CON PARTICOLARE ATTENZIONE ALLA DINAMICA CHE SI VERIFICA DOPO L'INIZIO DELLA TERAPIA IPOLIPEMIZZANTE;
- LE CURVE SONO RIPORTATE ALLE SLIDE CHE SEGUONO. I PAZIENTI SONO STATI RAGGRUPPATI PER RANGE DI LDL-C A INIZIO TERAPIA (CON L'OBIETTIVO DI VERIFICARE EVENTUALI COMPORTAMENTI DIVERSI A SECONDA DEL LIVELLO DI LDL-C DI PARTENZA).

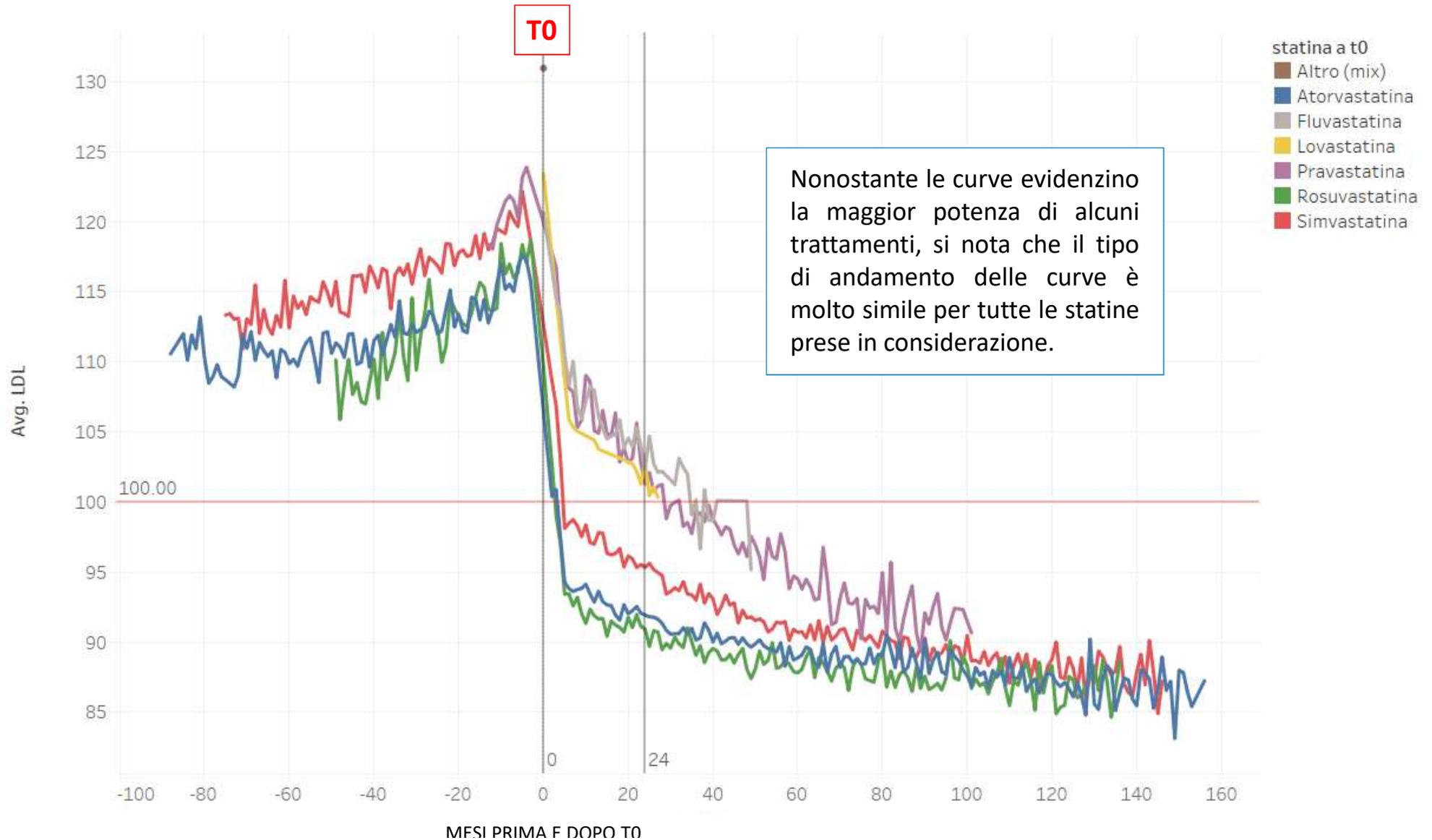
## TRACCIAMENTO DELL'ANDAMENTO DELL'LDL-C NEL TEMPO

## ANDAMENTO MEDIO DEI PAZIENTI PER RANGE DI LDL-C PRIMA E DOPO T0



Queste curve di andamento dell'LDL-C evidenziano come la maggior parte della riduzione di LDL-C ottenibile si verifichi nei primi 6 mesi di terapia.

# ANDAMENTO MEDIO DEI PAZIENTI PRIMA E DOPO T0, PER TIPO DI STATINA



# ANDAMENTO MEDIO DEI PAZIENTI PRIMA E DOPO T0, PER TIPO DI STATINA E RANGE DI LDL-C

**T0**



## L'ANALISI EVIDENZIA 4 TIPOLOGIE DI PAZIENTI NEL DATABASE (IN RELAZIONE ALL'INIZIO DELLA TERAPIA IPOLIPEMIZZANTE)

L'analisi delle curve ha evidenziato la presenza nel database di 4 tipologie di pazienti:

1. INIZIO TERAPIA IPOLIPEMIZZANTE CERTO. Pazienti per cui sono vere queste due condizioni:
  - la prima visita in cui è prescritto il farmaco ipolipemizzante è preceduta da visite in cui il farmaco non è presente;
  - il calo dell'LDL-C inizia dopo T0.

Per questi pazienti si assume che T0 sia l'effettivo inizio della terapia con ipolipemizzante).

2. INIZIO TERAPIA IPOLIPEMIZZANTE INCERTO PER PRIMA VISITA. Pazienti che, alla prima visita presente nel database, hanno già la prescrizione del farmaco ipolipemizzante: non è certo se quei pazienti abbiano iniziato a prendere il farmaco in quel momento o se siano arrivati al centro di diabetologia quando già prendevano il farmaco.

3. PROBABILE INIZIO TERAPIA IPOLIPEMIZZANTE PRIMA DI T0. Pazienti per cui:

- la prima visita in cui è prescritto il farmaco ipolipemizzante è preceduta da visite in cui il farmaco non è presente;
- il calo dell'LDL-C inizia dalla visita precedente a T0.

Per questi pazienti si assume che l'effettivo inizio della terapia con ipolipemizzante sia avvenuto prima di T0 (es. prescrizione fatta da altro specialista o dal MMG e non dal diabetologo).

4. INIZIO TERAPIA IPOLIPEMIZZANTE INCERTO PER MANCANZA di LDL-C ALLA VISITA PRECEDENTE A T0. Pazienti per cui:

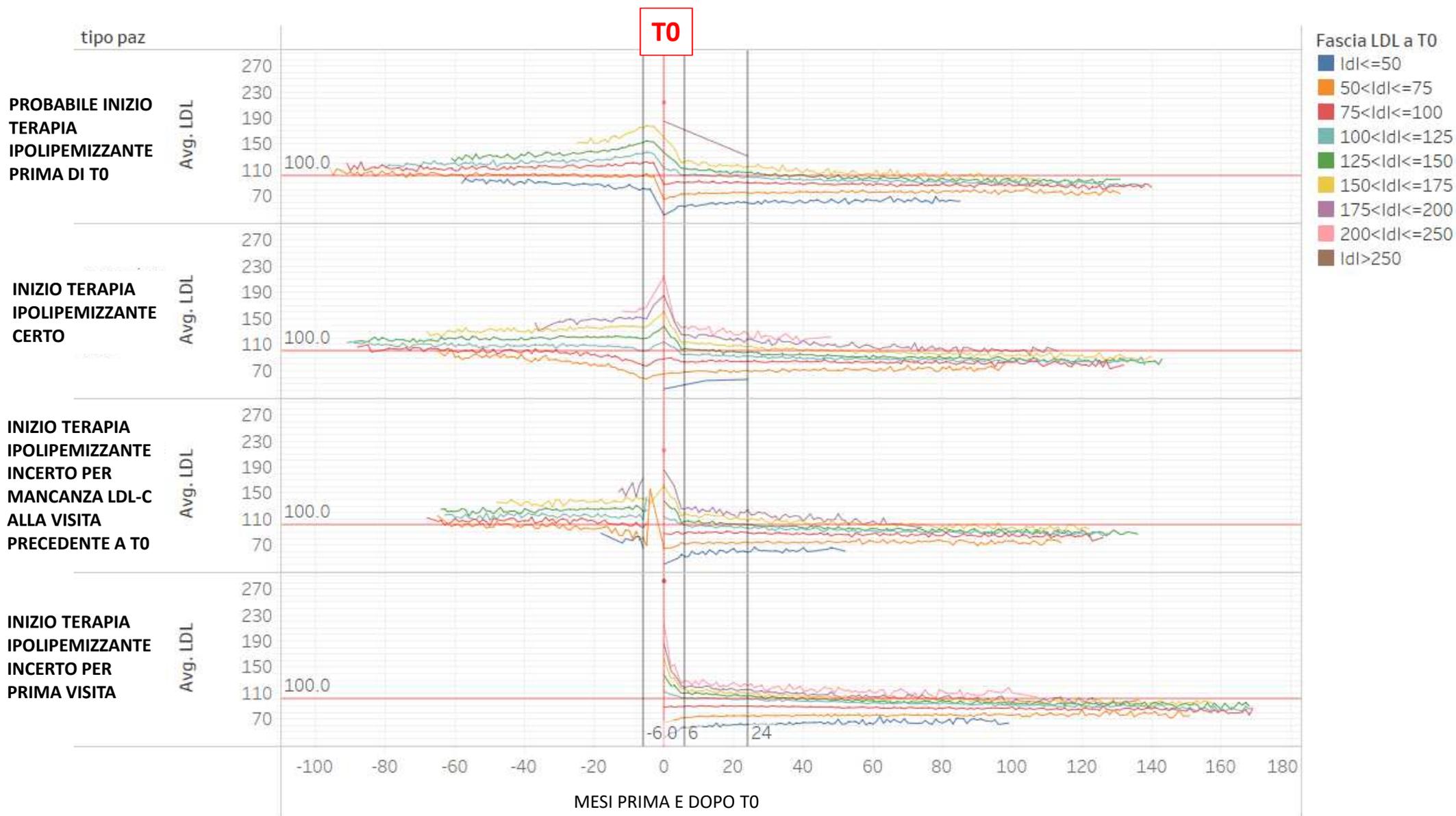
- la prima visita in cui è prescritto il farmaco ipolipemizzante è preceduta da visite in cui il farmaco non è presente;
- il valore dell'LDL-C alla visita precedente a T0 non è indicato nel database.

Per questi pazienti l'effettivo inizio della terapia con ipolipemizzante è incerto, in quanto potrebbero avere iniziato una terapia con ipolipemizzante prima di T0 (es. prescrizione fatta da altro specialista o dal MMG e non dal diabetologo) e avere già diminuito l'LDL-C a T0.

## TIPOLOGIE DI PAZIENTI E NUMEROSITÀ

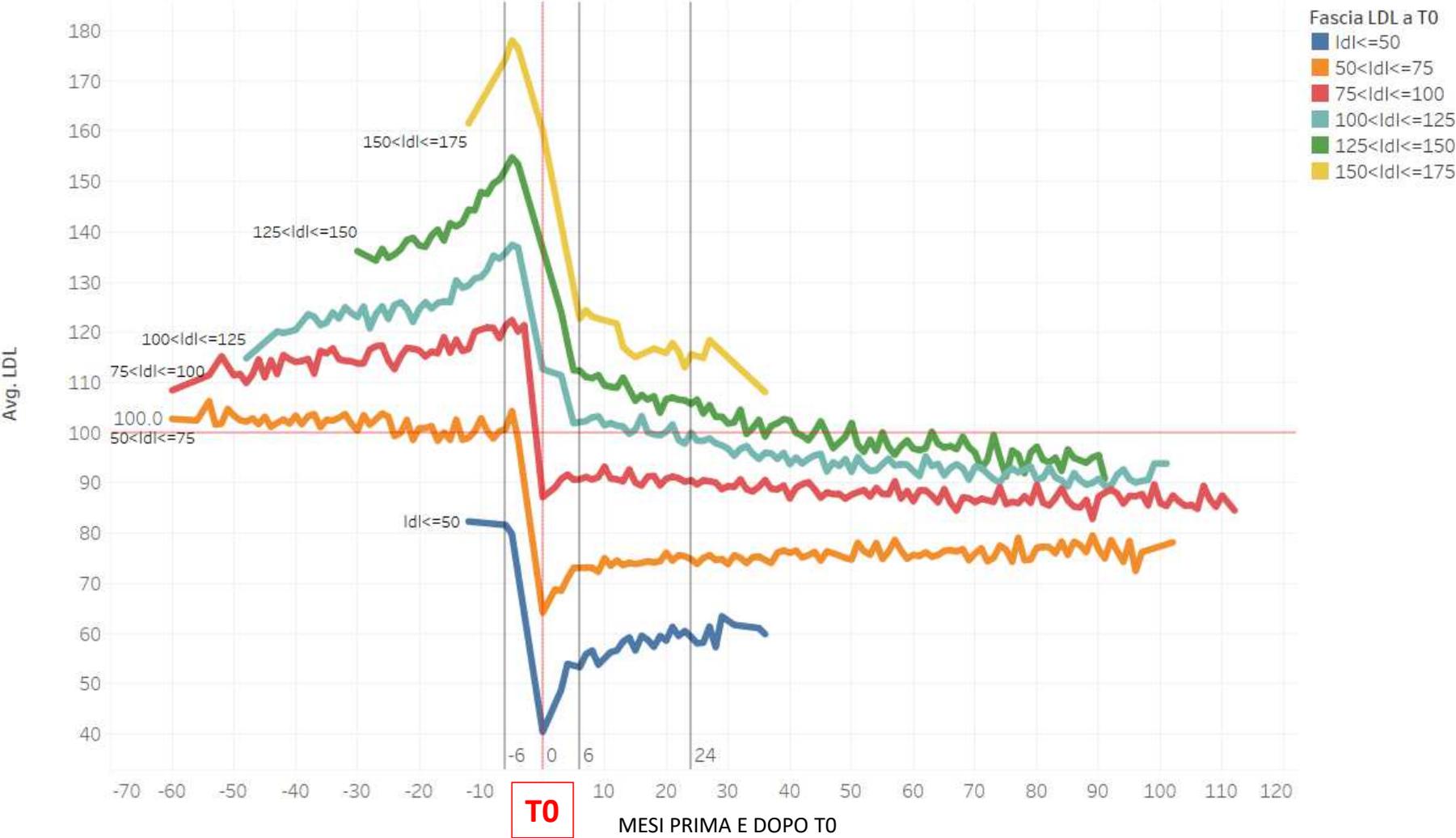
Tipo di paziente	Numerosità
INIZIO TERAPIA IPOLIPEMIZZANTE CERTO	37.415
INIZIO TERAPIA IPOLIPEMIZZANTE INCERTO PER PRIMA VISITA	72.416
PROBABILE INIZIO TERAPIA IPOLIPEMIZZANTE PRIMA DI T0	34.596
INIZIO TERAPIA IPOLIPEMIZZANTE INCERTO PER MANCANZA LDL-C ALLA VISITA PRECEDENTE A T0	24.586
TOTALE	169.013

# ANDAMENTO MEDIO DELLE 4 TIPOLOGIE DI PAZIENTI PER RANGE DI LDL-C PRIMA E DOPO T0

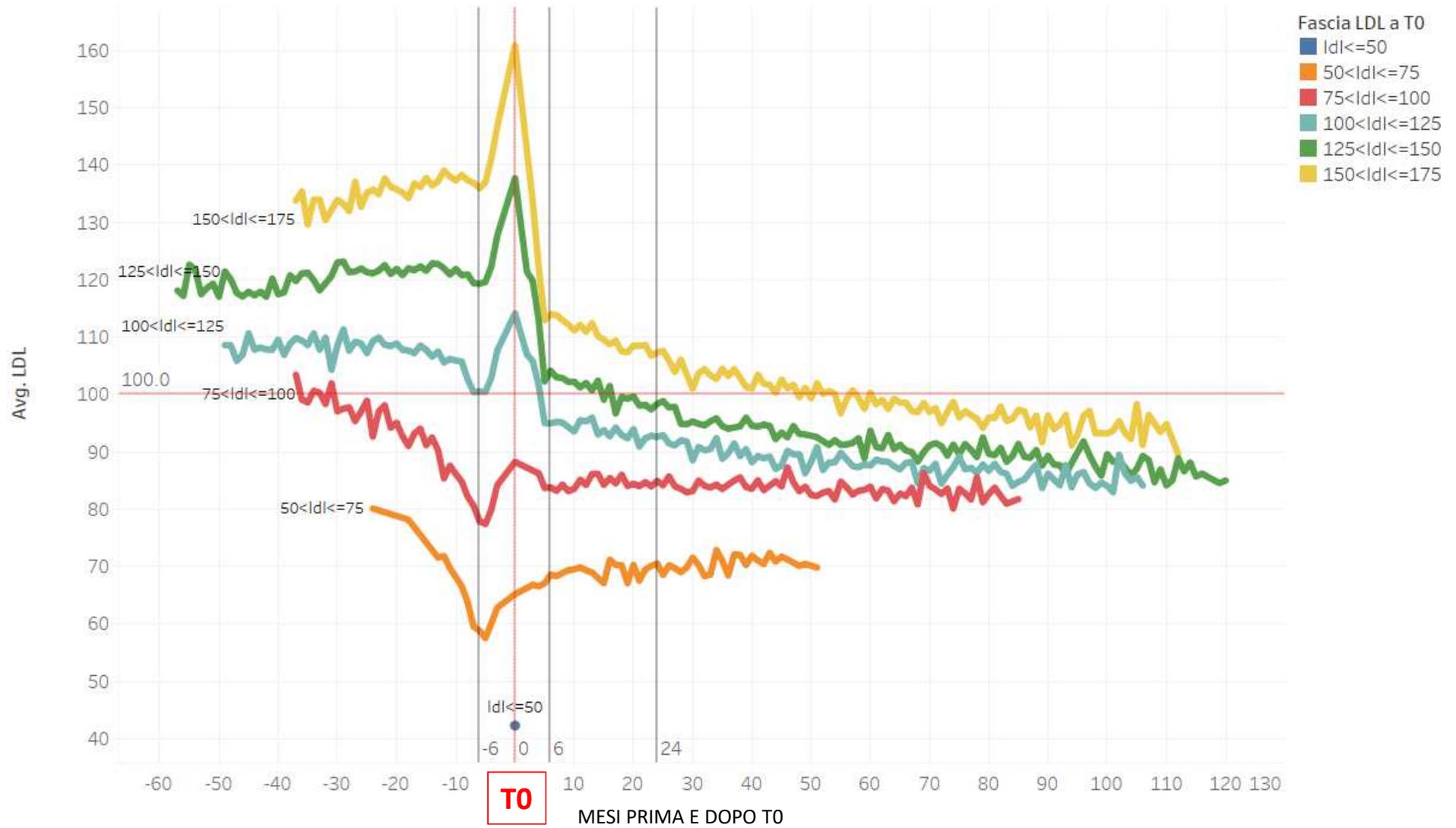


ANDAMENTO DELL'LDL-C NEL TEMPO: ZOOM SULLE FIGURE  
DELLE 4 TIPOLOGIE DI PAZIENTI DELLA SLIDE PRECEDENTE

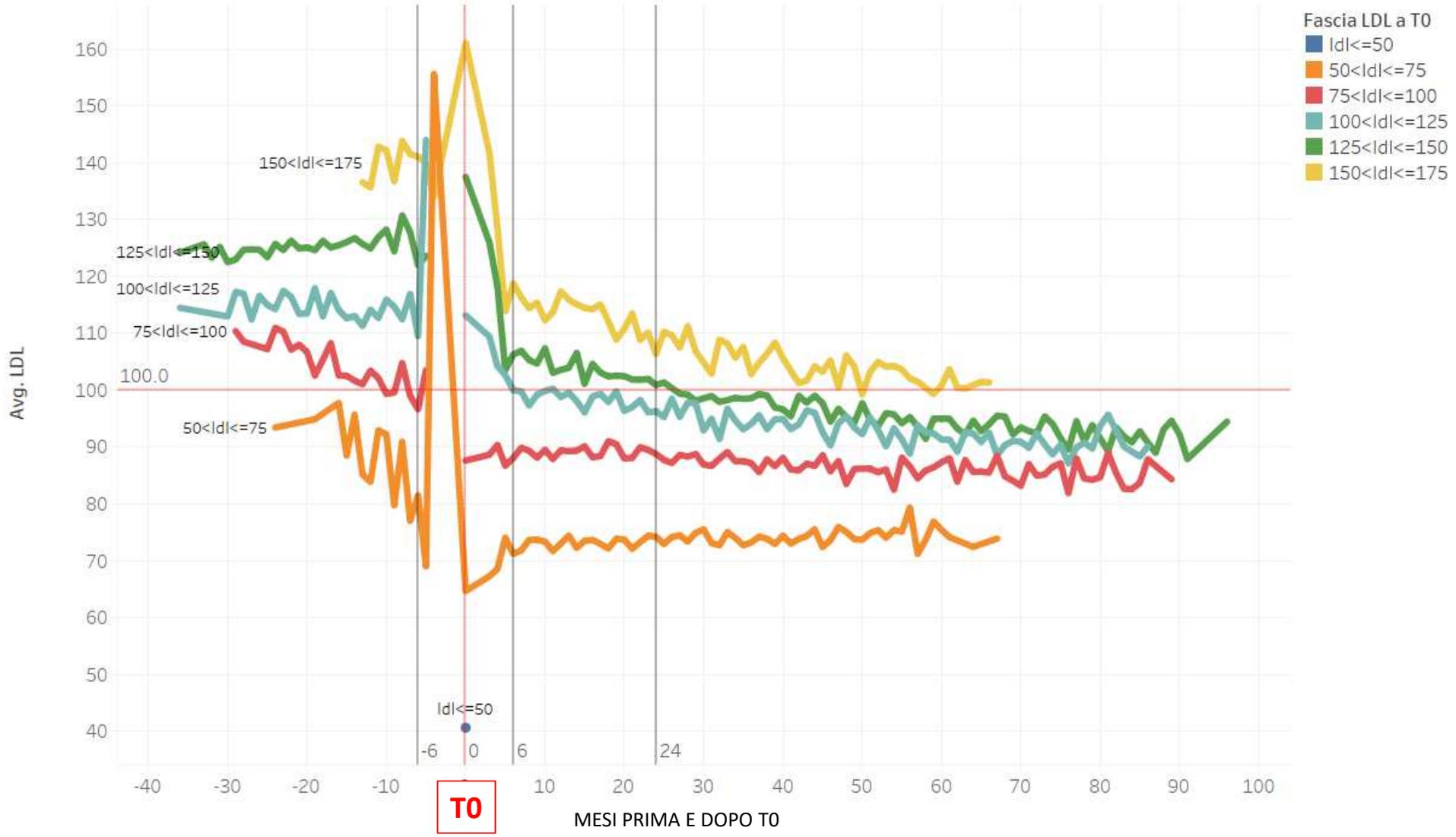
# ANDAMENTO MEDIO PAZIENTI CON PROBABILE INIZIO TERAPIA IPOLIPEMIZZANTE PRIMA DI T0 (PER RANGE DI LDL-C PRIMA E DOPO T0)



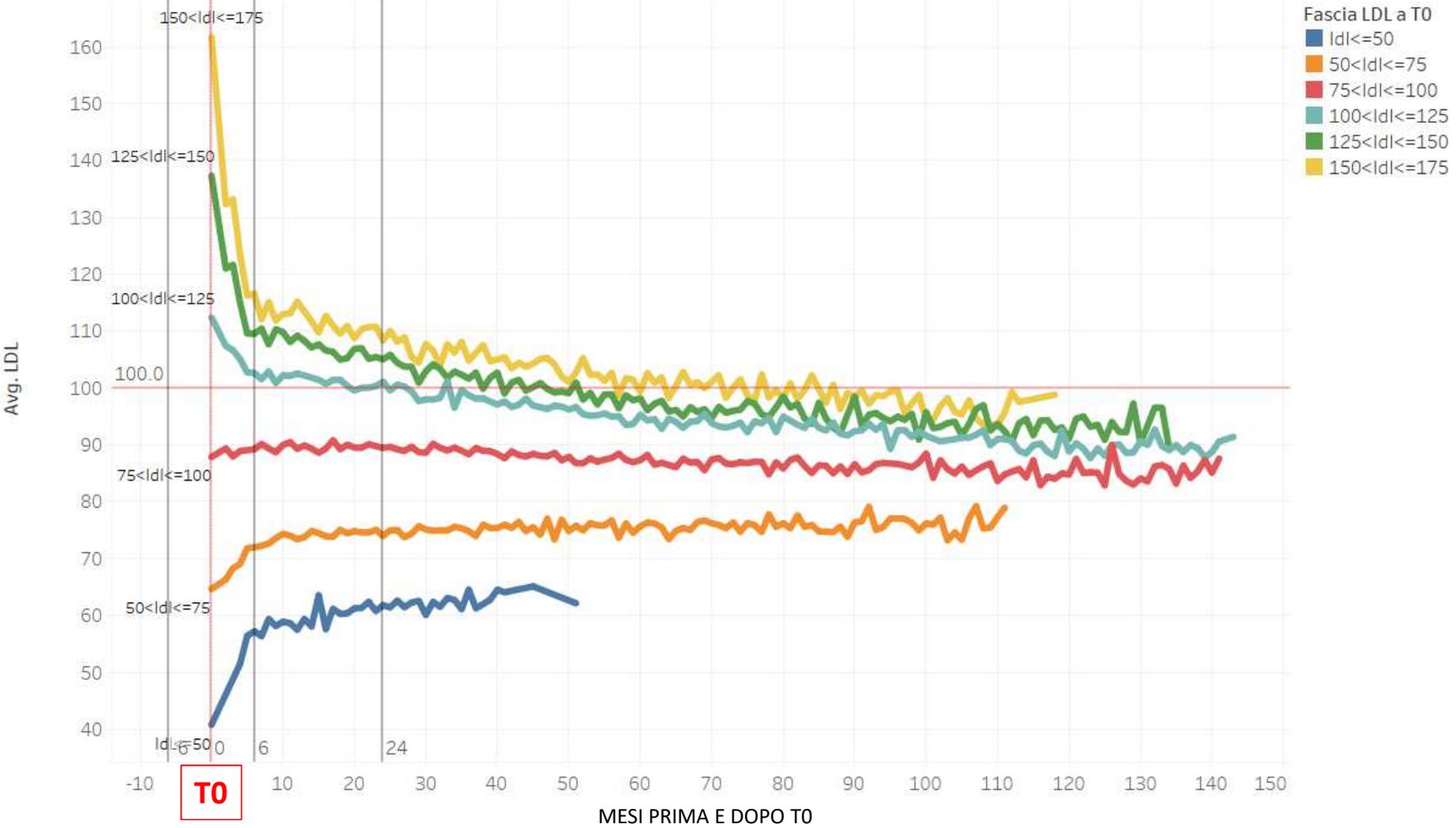
# ANDAMENTO MEDIO PAZIENTI CON INIZIO TERAPIA IPOLIPEMIZZANTE CERTO (PER RANGE DI LDL-C PRIMA E DOPO T0)



# ANDAMENTO MEDIO PAZIENTI CON INIZIO TERAPIA IPOLIPEMIZZANTE INCERTO PER MANCANZA DATO DI LDL-C ALLA VISITA PRECEDENTE A T0 (PER RANGE DI LDL-C PRIMA E DOPO T0)



# ANDAMENTO MEDIO PAZIENTI INIZIO TERAPIA IPOLIPEMIZZANTE INCERTO PER PRIMA VISITA (PER RANGE DI LDL-C PRIMA E DOPO T0)



## ANALISI MACHINE LEARNING: APPROFONDIMENTI SECONDA FASE

IL RUOLO CHIAVE RAPPRESENTATO DAL GRADO DI RIDUZIONE DELL'LDL-C ALLA VISITA SUCCESSIVA A T0:  
RICERCA DEI VALORI SOGLIA CORRELATI AL RAGGIUNGIMENTO DEL TARGET

## ANALISI MACHINE LEARNING: APPROFONDIMENTI SECONDA FASE

Dalla coorte sono stati selezionati i pazienti **con inizio terapia certo** (37.415 individui) per migliorare l'ulteriore apprendimento dell'algoritmo, eseguito solo sui pazienti con 'le caratteristiche più pulite'.

Dopo che il machine learning della prima fase ha evidenziato il ruolo dominante (rispetto a tutte le altre caratteristiche del paziente) del rapporto fra LDL-C a inizio terapia e LDL-C alla visita successiva, gli approfondimenti hanno avuto il seguente obiettivo:

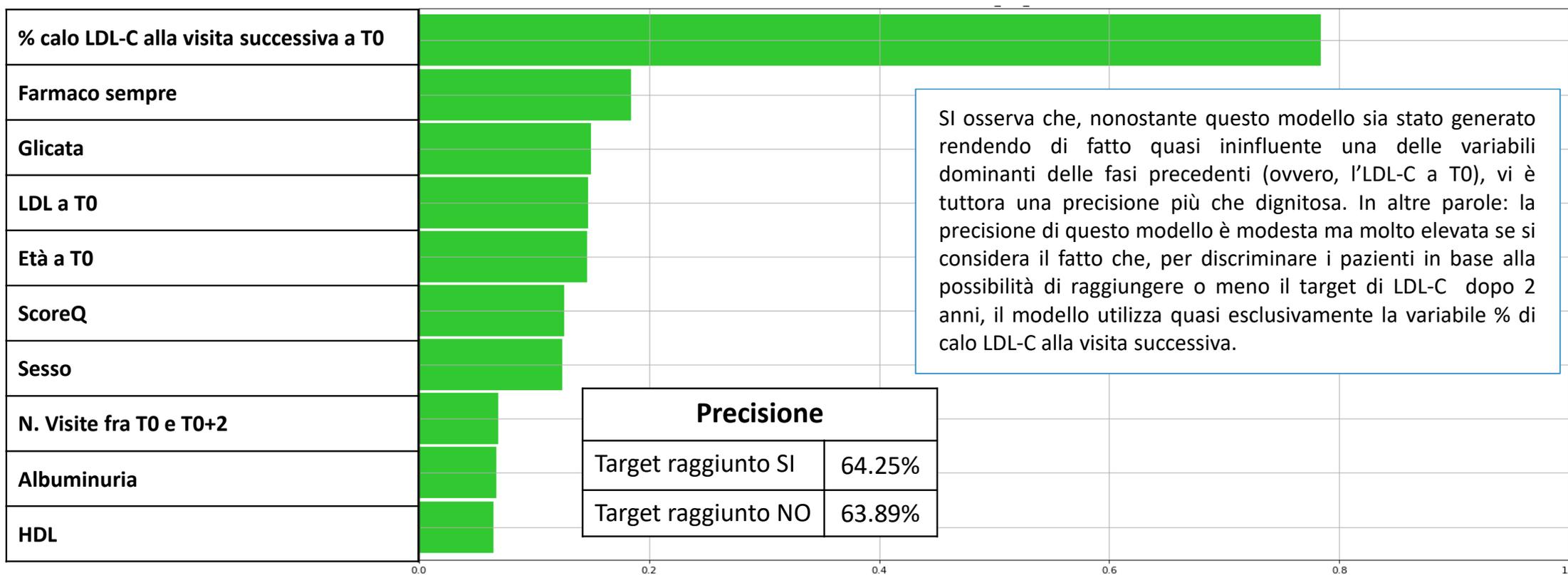
**La ricerca, per diversi range di LDL-C a inizio terapia ( $>100 \leq 125$ ,  $>125 \leq 150$ ,  $>150 \leq 175$ ,  $>175 \leq 200$ ,  $>200 \leq 250$ ), della riduzione di LDL-C minima necessaria correlata ad una maggiore probabilità di raggiungere il target dopo 2 anni da inizio terapia.**

**Il machine learning ha prodotto, per ognuno dei range di LDL-C indicato, la riduzione minima che deve essere ottenuta entro la visita successiva per poter (probabilmente) raggiungere il target dopo 2 anni.**

**Alla slide che segue vengono riportati, a titolo esemplificativo, alcuni dettagli sui risultati di uno dei modelli generati (per il range di LDL-C  $>100 \leq 125$ ).**

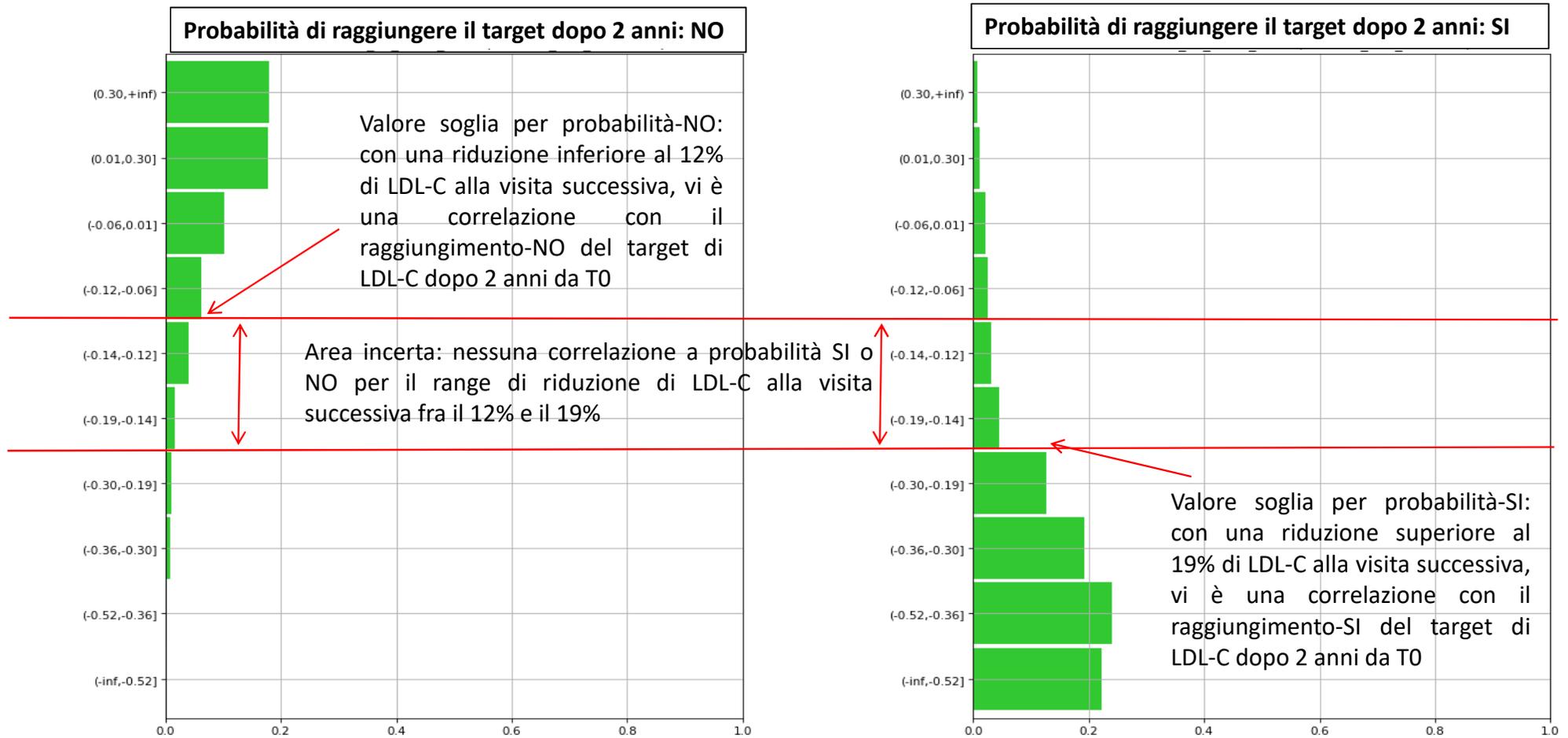
## MACHINE LEARNING: RICERCA SOGLIA MINIMA DI RIDUZIONE LDL-C ALLA VISITA SUCCESSIVA PER RANGE DI LDL-C A T0 >100<=125

Il risultato evidenzia il ruolo dominante della % di calo dell'LDL-C alla visita successiva rispetto alla probabilità di raggiungere il target dopo 2 anni da T0.  
 Avendo filtrato il campione per range di LDL-C, il **valore di LDL-C di partenza gioca un ruolo quasi irrilevante, in quanto tutti i pazienti di questo gruppo hanno livelli di LDL-C di partenza molto vicini.**



La lunghezza delle barrette verdi si riferisce all'importanza relativa fra le variabili (Thorsen-Meyer; Hans-Christian; Nielsen, Annelaura B. et al / Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records. In: The Lancet Digital Health. 2020 ; Vol. 2, No. 4. pp. e179-91.)

# I VALORI SOGLIA PROPOSTI DAL MACHINE LEARNING PER QUANTO RIGUARDA LA % DI RIDUZIONE DI LDL-C ALLA VISITA SUCCESSIVA CORRELATA ALLA MAGGIORE O MINORE PROBABILITÀ DI RAGGIUNGERE IL TARGET DOPO 2 ANNI (PER PAZIENTI CON RANGE DI LDL-C A T0 >100<=125)



## RISULTATI MACHINE LEARNING: I VALORI SOGLIA PER I DIVERSI RANGE DI LDL-C

La tabella riporta i valori soglia indicati dal machine learning per i **diversi range di LDL-C a inizio terapia (>100&lt;=125, >125&lt;=150, >150&lt;=175, >175&lt;=200, >200&lt;=250)**, relativi alla % di riduzione di LDL-C correlata alla probabilità di raggiungere (o meno) il target di LDL-C <100 dopo 2 anni da inizio terapia.

Range di LDL-C a inizio terapia (T0)	Valori soglia indicati dai risultati del machine learning	
	Probabile target LDL-C dopo 2 anni: SI	Probabile target LDL-C dopo 2 anni: NO
>100&lt;=125	Riduzione LDL-C alla visita successiva a T0 >19%	Riduzione LDL-C alla visita successiva a T0 <12%
>125&lt;=150	Riduzione LDL-C alla visita successiva a T0 >29%	Riduzione LDL-C alla visita successiva a T0 <29%
>150&lt;=175	Riduzione LDL-C alla visita successiva a T0 >36%	Riduzione LDL-C alla visita successiva a T0 <31%
>175&lt;=200	Riduzione LDL-C alla visita successiva a T0 >49%	Riduzione LDL-C alla visita successiva a T0 <42%
>200&lt;=250	Riduzione LDL-C alla visita successiva a T0 >48%	Riduzione LDL-C alla visita successiva a T0 <41%

## ANALISI MACHINE LEARNING: VERIFICA RISULTATI SUI VALORI SOGLIA

In base alle indicazioni fornite dal machine learning relativamente ai valori soglia sulle % di calo di LDL-C alla visita successiva, è stata fatta una fotografia di verifica (vedi slide successiva) sulle % di riduzioni medie alla visita successiva, per range di LDL-C a T0.

La fotografia è stata fatta per tutte le 4 tipologie di pazienti descritte alla slide 26.

I valori soglia indicati dal machine learning sono coerenti con i risultati riportati nella fotografia di verifica.

# VARIAZ. % MEDIE DI LDL-C ALLA VISITA SUCCESSIVA A T0 CORRELATE ALLA PROBABILITÀ DI RAGGIUNGERE (O MENO) IL TARGET A T0+2ANNI



FINE DEL REPORT

Report realizzato con il contributo non condizionante di Daiichi Sankyo Italia S.p.A.

