Enabling FHIR for Clinical Research: Standardizing the FHIR Python API

Zexi (Amber) Zhou, Harold R. Solbrig, Christopher G. Chute
Johns Hopkins Institute for Clinical and Translational Research
Johns Hopkins University

Abstract

- FHIR (Fast Healthcare Interoperability Resources) is a standard for health care data exchange, published by HL7 (Health Level Seven International).
- Clinical research incorporates big data to support practice quality improvement and translational research applications.
- The use of a FHIR Python API (Application Programming Interface) allows the manipulation and transformation of FHIR resources.
- Current Python API (tooling) needs to be enhanced and documented for wider use and uptake.

Objectives

This project aims to
1) demonstrate how a FHIR Python API can be used to simplify FHIR derived data analysis and make it more readable and easier to conduct
2) show efforts to enhance and document the Python API for wider use and uptake

Use of fhirclient:
Load and Format Data from a FHIR Server

```python
import fhirclient.models.patient as p

# search patients with address-country and flatten JSON into table
search = p.Patient.where(struct={'address-country': 'USA'})
patients = search.perform_resources(smart.server)

flat = flatten_json(patients[0].as_json())
result = json_normalize(flat)

for patient in patients:
    f = flatten_json(patient.as_json())
    r = json_normalize(f)
    result = pd.concat([result, r], axis = 0, ignore_index = True, sort = True)

result = result.reset_index(drop=True)
result = result.drop(0)
result = result.replace(np.nan, '', regex = True)
```
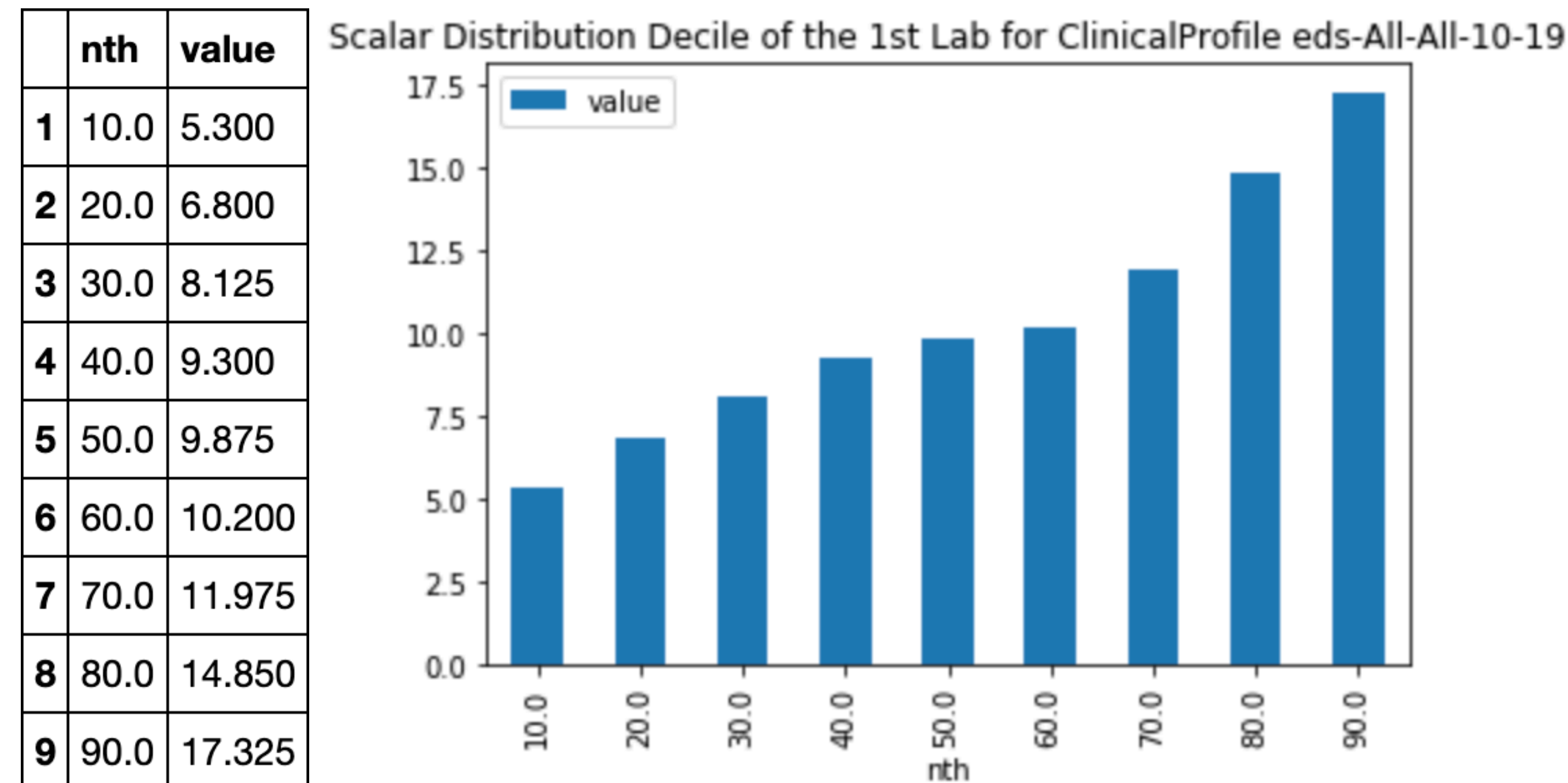
	id	identifier_value	active	name_given	name_family	name_use	telecom_system	telecom_use	telecom_value
1	60776	NCC-1032			Bajwa		phone	home	+1(555) 555-1212
2	60838	NCC-1627	True	Batman	Kasi		phone	home	+1(555)999-3232
3	75848	11396810			FREDERICKS	usual		home	856-555-1212
4	1376774			gm_test	test		phone	home	1234567890
5	1437875	454721		JOHN	DOE	usual	phone	home	(216)123-4567
6	1437876	454721		JOHN	DOE	usual	phone	home	(216)123-4567
7	1437884			John	Doe				
8	1437934	12001		John	Jones				
9	1440598	12001		John	Jones				
10	1440820			John	Doe				

Visualize Clinical Profile Examples

	nth	value
1	10.0	5.300
2	20.0	6.800
3	30.0	8.125
4	40.0	9.300
5	50.0	9.875
6	60.0	10.200
7	70.0	11.975
8	80.0	14.850
9	90.0	17.325

Scalar Distribution Decile of the 1st Lab for ClinicalProfile eds-All-All-10-19

eds-All-All-10-19: Ehlers-Danlos syndrome, all gender, all race, age 10-19
eds-All-All-All: Ehlers-Danlos syndrome, all gender, all race, all age

lab_name	lab_code	max_base	min_base	median_base	mean_base	stdDev_base	max	min	median
Urea nitrogen [Mass/volume] in Venous blood	12962-7	70.00	1.00	11.000	11.976	0.868	33.00	2.00	9.500
Alanine aminotransferase [Enzymatic activity/v...	1742-6	896.00	4.00	17.500	28.116	27.990	180.00	6.00	15.000
Calcium [Mass/volume] in Serum or Plasma	17861-6	11.40	4.90	9.100	9.010	0.144	10.70	7.20	9.250
Magnesium [Mass/volume] in Serum or Plasma	19123-9	8.40	1.20	2.000	2.083	0.175	2.50	1.60	2.000
Aspartate aminotransferase [Enzymatic activity...	1920-8	1203.00	6.00	19.000	26.077	32.782	70.00	7.00	18.000
Bilirubin.total [Mass/volume] in Serum or Plasma	1975-2	4.20	0.10	0.400	0.436	0.109	2.30	0.20	0.325
Cholesterol in HDL [Mass/volume] in Serum or P...	2085-9	109.00	18.00	58.000	57.957	1.441	74.00	40.00	50.500
Cholesterol [Mass/volume] in Serum or Plasma	2093-3	349.00	74.00	178.750	185.627	3.630	254.00	97.00	171.000

Comparison of Scalar Distribution of Labs for ClinicalProfiles eds-All-All-10-19 (gray) and eds-All-All-All (blue)
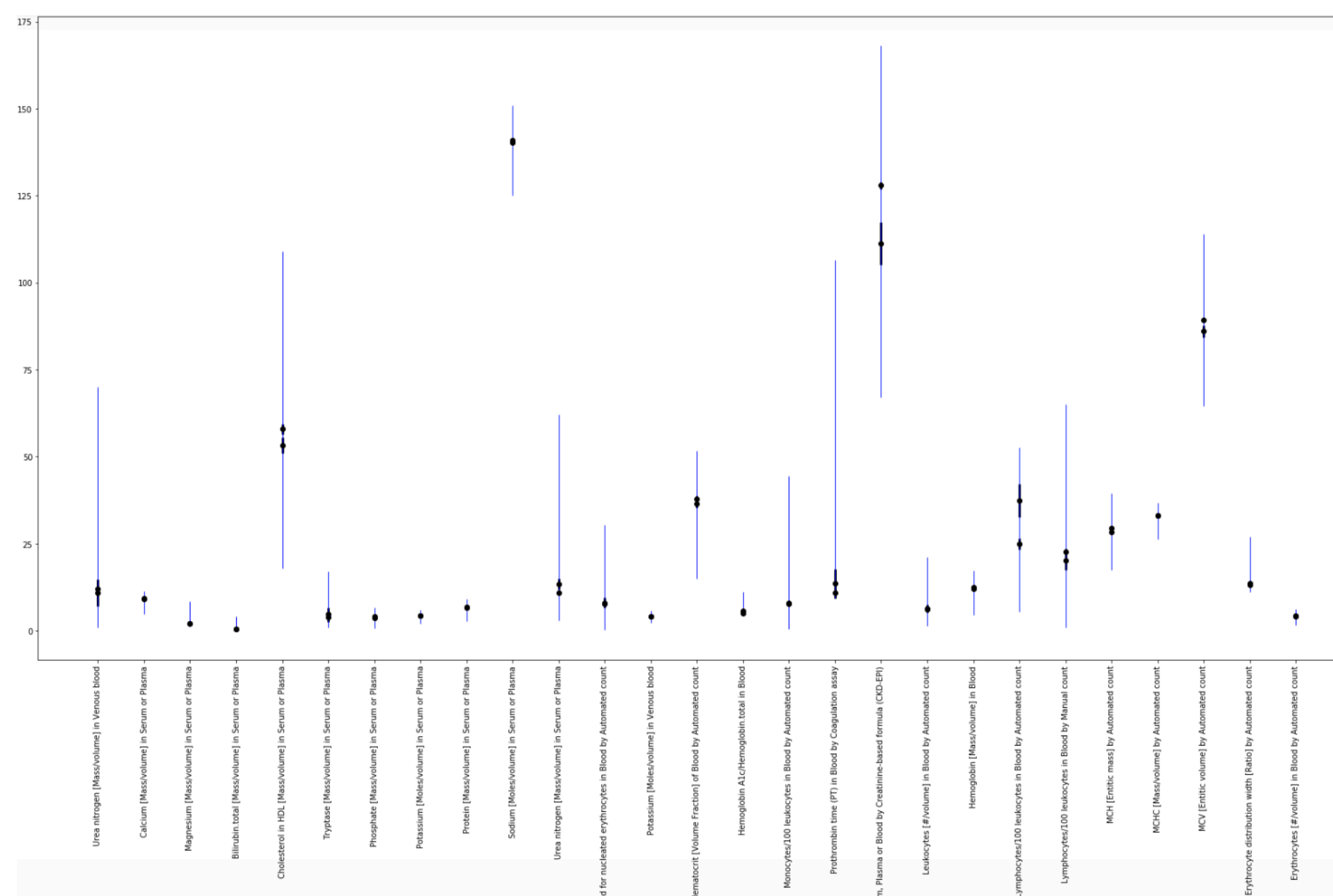
Client-py Enhancements

provide type hints using dataclass and typing library in template-resource for fhir model generation

```python
@dataclass
class Medication(domainresource.DomainResource):
    """ Definition of a Medication.

    This resource is primarily used for the identification and definition of a
    medication for the purposes of prescribing, dispensing, and administering a
    medication as well as for making statements about medication use.
    """
    resource_type: ClassVar[str] = "Medication"
    amount: Optional[ratio.Ratio] = None
    batch: Optional[MedicationBatch] = None
    code: Optional[codeableconcept.CodeableConcept] = None
    form: Optional[codeableconcept.CodeableConcept] = None
    identifier: Optional[List[identifier.Identifier]] = empty_list()
    ingredient: Optional[List[MedicationIngredient]] = empty_list()
    manufacturer: Optional[fhirreference.FHIRReference] = None
    status: Optional[str] = None

    jsondict: InitVar[Optional[dict]] = None
    strict: InitVar[bool] = True

    def __post_init__(self, jsondict, strict) -> None:
        fhirabstractbase.FHIRAbstractBase(jsondict, strict)
```

Next steps to implement include:
set constraints, size, value ranges for primitive data types; shortcuts; pypi; extensions; documentation, etc

Github link: https://github.com/zzhou41/client-py

Conclusions

- The use of FHIR Python API enhances the automation of data science prep work for machine learning, replacing the traditional manual prep process.
- Developed tools help FHIR derived data analysis for clinical research.
- Ongoing improvements on client-py will encourage further use.

Acknowledgements

Contact Information:
zzhou41@jhu.edu
443-449-0131
Github: https://github.com/zzhou41