PharmaSUG 2019 - Paper PO-305

Merging Sensor Data with Patient Records in Clinical Trials – Systems and Benefits

Surabhi Dutta, EG Life Sciences, Somerset, New Jersey, Shubhranshu Dutta, Annandale, New Jersey

ABSTRACT

Patient care involves data capture in all phases of care delivery including clinical trials, provider visits, lab work, sensor data from wearable sensors and patient generated data from hand held devices. We are accustomed to devices that generate health indicator data in large volumes and rapid rate. This paper discusses the benefits, challenges and methods of merging these disparate datasets characterized by volume, variety, velocity.

The process of healthcare in the current scenario is a one-way system where patient health indicator data is entered in the Electronic Health Records by a healthcare practitioner. This one sided approach is generally based on the health conditions at the date and time of patient and provider contact. While the purpose of Electronic Health Records is recording patient data during a visit to the provider, patient reported outcomes are often ignored in the process. An electronic health record (EHR) contains patient health information that can be shared across healthcare providers in multiple platforms. Information captured includes: Administrative and Billing data, Patient demographics, Progress notes, Vital signs, Medical histories, Diagnoses, Medications, Immunization dates, Allergies, Radiology images, Lab and test results.

INTRODUCTION

With technological advancement and patient health awareness, telemedicine and wearable sensors many chronic conditions can be managed with precision and doctors can get a holistic view of patient health. Patient generated data from wearable sensors capture a variety of health indicators and these are not limited to: Mortality indicators, Morbidity indicators, Health status Disability indicators, Nutritional indicators, Social and mental health indicators, Health system indicators and Health determinants.

Sensor data for self-reported health indicators may be used for a variety of purposes not limited to: recommendations, alerts, self-assessment, patients may choose to classify their current health status as good, fair, or poor, daily activity tracking, severity of acute and chronic morbidity conditions, time of medications and benefits observed, drug interactions including patient inputs for any adverse events that could be related to drug interactions and prior or concomitant medications, physical disability based on observed activities of daily living, current nutritional status as measured by body mass index and real-time monitoring of patient vitals.

APPLICATION AND SOLUTIONS

a) Types of wearable devices: There are a variety of wearable devices these include: fitness trackers, smart watches, connected headsets, smart glasses, wrist bands. One example is the Continuous glucose monitoring (CGM) device that helps to monitor glucose levels, so that patients can see patterns and trends to help better manage diabetes and practice self-tracking.

Table 1 Examples of wearable sensors

Device type	Data collected	Examples
Wrist worn	Actigraphy, HR (Heart Rate), BP (Blood Pressure), EDA (Electrodermal activity)	Actiwatch Spectrum by Phillips, ActiGraph Link by ActiGraph, E4 by Empatica, ViSi Mobile by Sotera Wireless
Skin patch	ECG (Electrocardiography), actigraphy, skin temperature	BioStampRC by MC10, HealthPatch by Vital Connect, BodyGuardian by Preventice
Cuffs	BP, HR	Intellisense Digital BP Monitor by Omron Healthcare
Finger worn	HR, SpO2	iSpO2 Pulse Oximeter by Massimo
Clothing embedded sensors	HR, HRV (Heart Rate Variability), ECG, Breathing Rate, actigraphy	Smart shirts by Hexoskin
Headbands	EEG (Electroencephalogram), EMG (Electromyography)	EMOTIV EPOC by Emotiv, 4D FORCE by 4D FORCE

Image source: https://www.researchgate.net/figure/Examples-of-wearable-sensors_tbl1_321578511
Figure 1. Examples of wearable sensors

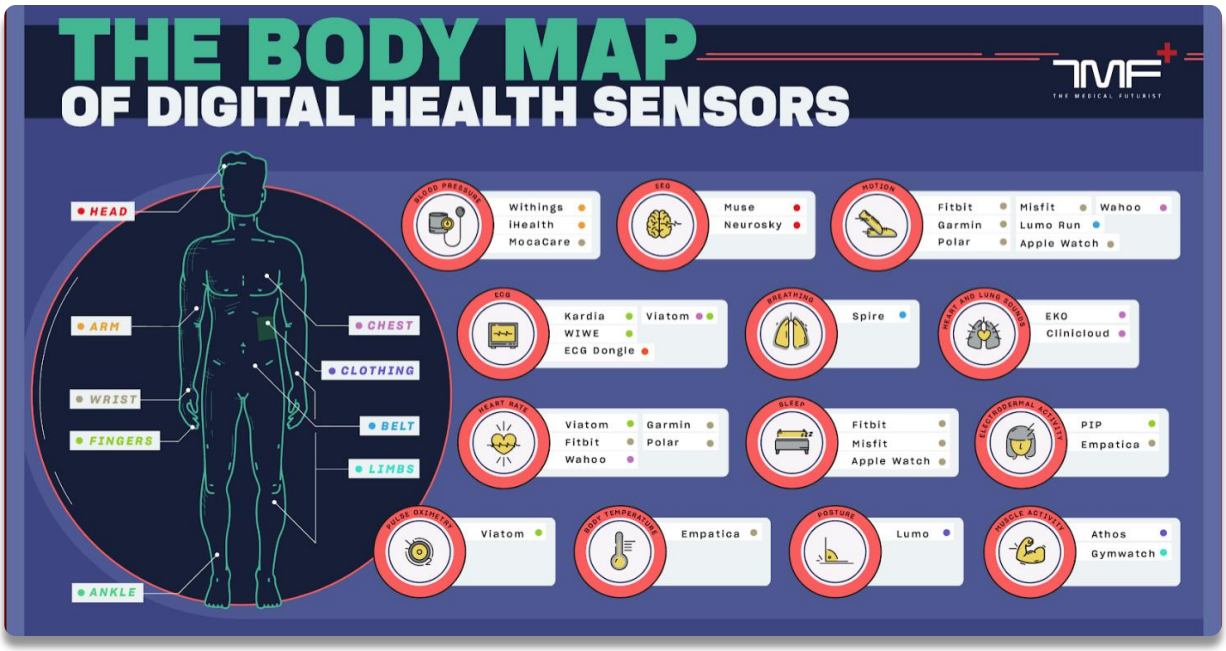


Image source: https://medicalfuturist.com/the-body-map-of-digital-health-sensors
Figure 2. Wearable Sensors Usage

b) Utilization of data generated by sensors: Sensors can be personal devices, wearable or handheld that have been calibrated and checked for accuracy at the Physician's office. The access to data could be direct (we would need to develop and install native apps both in the wearable and in the smartphone) or indirect (where data is collected in a warehouse and it can be retrieved by a server on an ongoing or daily basis). Most Wearable devices facilitate access to raw data and allow the creation of different algorithms via research tools (e.g. Matlab®, R) to achieve insightful patient assessment and promote evidence based decision making. Fitbit data can be downloaded as a flat file (excel sheet).

c) Why we need to merge these data (Purpose and Benefits):
We need to merge these data to generate a 360 degrees, comprehensive view of the patient's health.
A comprehensive view would be an overall summary where signals are stored as complex/raw time series data by the data acquisition electronics.

Table 2 Novel endpoints: application, benefits, and examples

Application	Benefit	Examples and references
Safety monitoring/patient phenotyping	• Early safety signal, dose and frequency adjustments, discontinuation of certain drug candidates • Better understanding of mechanistic and pharmacological drug profile if combined with PK and wet lab test data	Vital sign, e.g. HR, RR, skin temperature, BP, and actigraphy[37,39]
Novel endpoints	• Mobility as a measure of quality of life • Sleep studies in the home settings for extended periods of time • More sensitive measures than traditional clinical scales in movement disorders	• Actigraphy in Oncology[53] • Actigraphy as a measure sleep in a home settings[54–56] • Gait and tremor in Parkinson's disease[57,58]
Medication adherence monitoring and intervention	• Improved adherence • Informed decisions about dose adjustments • Increased efficiency in postmarket data collection	• Adherence surveys • Drug intake reminder apps • Objective data on drug intake - smart cap bottles
Patient enrollment and retention in clinical trials	• Fewer obstacles to enroll in clinical trials • Reduced burdens for patients to participate • Increased patient outreach	• Remote enrollment and consent apps • Reminder apps about study procedures and clinical trial progress

Image source: https://ascpt.onlinelibrary.wiley.com/doi/pdf/10.1002/cpt.966
Figure 3. Application and Benefits of Sensor Generated Data

- Monitoring patient medication adherence: Sensor data may be effectively used in monitoring any missed dose. This will boost adherence.
- Close tracking of adverse events and drug efficacy: Sensor data will be a very helpful tool to close track any adverse events and generate alerts for patients and caregivers during a trial period. Similarly, any health improvements and drug related benefits may be noted.
- Time point based patient vitals tracking: Time point based patient vitals data capture and tracking will establish better care and prevent serious adverse events.
- Concurrent medications and drug interaction monitoring: Capturing sensor data and merging with concomitant medications and patient history data can be an effective way to prevent drug related adverse events.
- Time Series Analysis: Patient generated data can be used for time series/trend analysis. This can be done with the help of various statistical analysis tools to predict and interpret results.
  Data presentations and visualizations can be generated at real time for adverse event monitoring.
- Data presentations could be included for some of the sensor, wearable devices captured and patient generated endpoints like Average Heart rate, Systolic Blood Pressure (SBP), Diastolic Blood Pressure (DBP), Sleep, Pain in a detailed Summary of Demographics and Baseline Characteristics table. Any adverse events captured by sensor data can also be included in the Adverse Events tables.
- Data standardization and harmonization can enable quality control and usability for analysis.

d) Challenges in merging these data:

- The sensor data structure is very different compared to traditional data collected at predefined time points by clinical sites. This data consists of multiple layers: raw unfiltered data, raw filtered data to eliminate invalid data in accordance with the scoring algorithms.
- Additionally, there are no well-developed standards that would help to organize, annotate, and standardize the data and provide data mapping tools to electronic data capture databases.
- Volume, velocity and variety: Clinical data is in standardized formats while sensor generated data is humongous.
- Constant patient monitoring with time series analysis on real time on smartphones versus integration of sensor data with trial data in standardized formats remains a challenge as smartphones have their own inbuilt analytics framework for analyzing sensor generated data. Trial data is highly structured and formatted for time point based analysis.
- Clinical data is in standardized format, merging patient generated/sensor data with clinical data will need time point based data integration. Data may be pooled from sensors at regular time intervals.
- Data validation and missing data: Sensor data may have issues related to data checks when it is not harmonized and some of the data points are missing.
- Multi-sensor data merging has many issues and obstacles, but pragmatic implementation remains equally challenging, e.g. access to commercial data, knowledge sharing between companies, patient safety, privacy and HIPPA laws.
- Acceptability and patient usage training: Patients need to devote time to get accustomed to using the wearable devices and understand the benefits of data monitoring.

SOLUTIONS AND METHODS OF MERGING SENSOR DATA WITH CLINICAL TRIAL DATA:

a) Data Standardization and Harmonization: Building a Common Data Model:

When using data from multiple systems and applications together, we need to aggregate common data elements that can save effort, streamline development, and help with faster analytics. Common Data Model is a set of standardized, extensible data composed of predefined schemas that includes entities, attributes, semantic metadata, and relationships. These schemas help map and harmonize data to simplify the creation, aggregation, and analysis.

One such example is the PCORnet program (Patient Centered Outcomes Research Institute (PCORI)). The Common data model by PCORI was developed based on used cases to accommodate various data sources and types. Patient generated data from wearable devices can be harmonized in to required variables using the Common Data Model and integrated into trial data.
Common Data Models can be described in terms of the elements and attributes as key-value pairs, as in a typical XML (eXtensible Markup Language) or JSON (JavaScript Object Notation) specification.

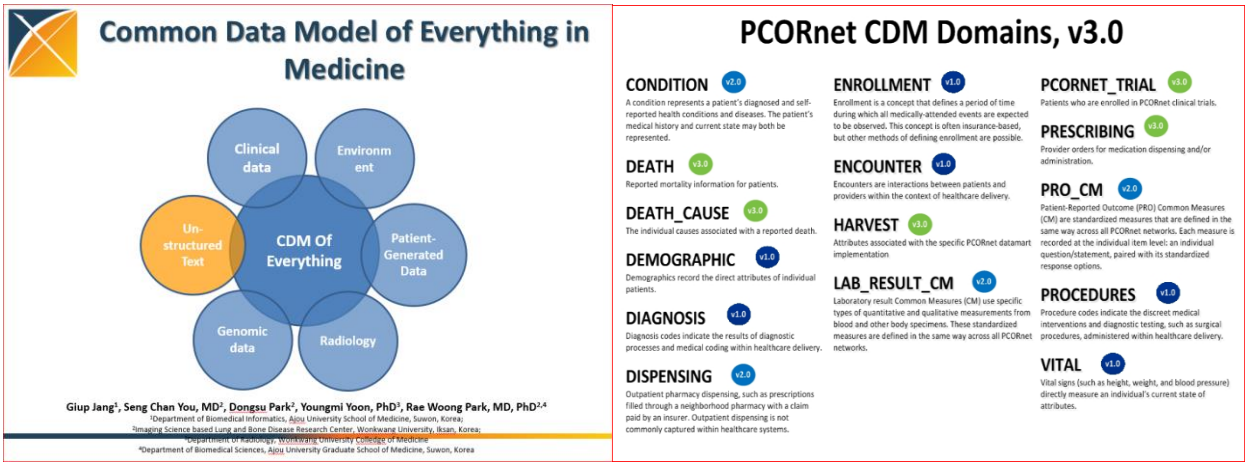


Image Source:
www.ohdsi.org/wpcontent/uploads/2015/04/OHDSI_dimesional_expansion_of_data18.8.14.pdf
Figure 4. Building a Common Data Model

b) Accessing sensor data: The data capture workflow using sensors can be visualized as:

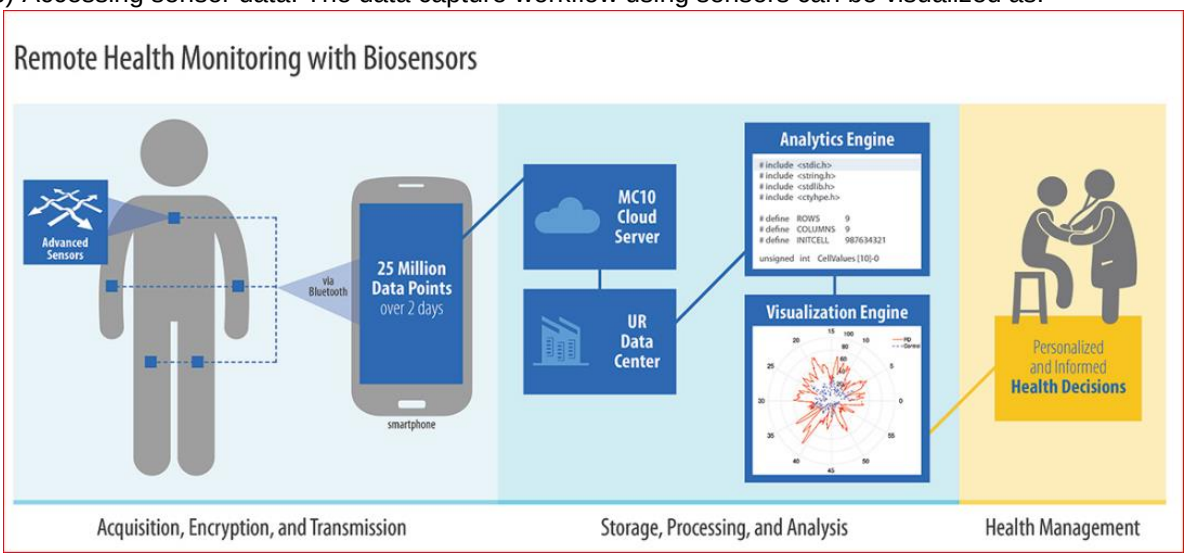


Image Source: https://everbetter.rochester.edu/2017/05/01/skin-sensors-provide-wealth-of-patient-data/
Figure 5. Data Capture Workflow

One example of patient data capture is the Medidata Patient Cloud: Patient reported cloud based outcomes and vitals data with Electronic data Capture. (Please refer the following link for Medidata Epro Patient generated data capture. https://www.youtube.com/watch?time_continue=5&v=3B4k3n8nriM).

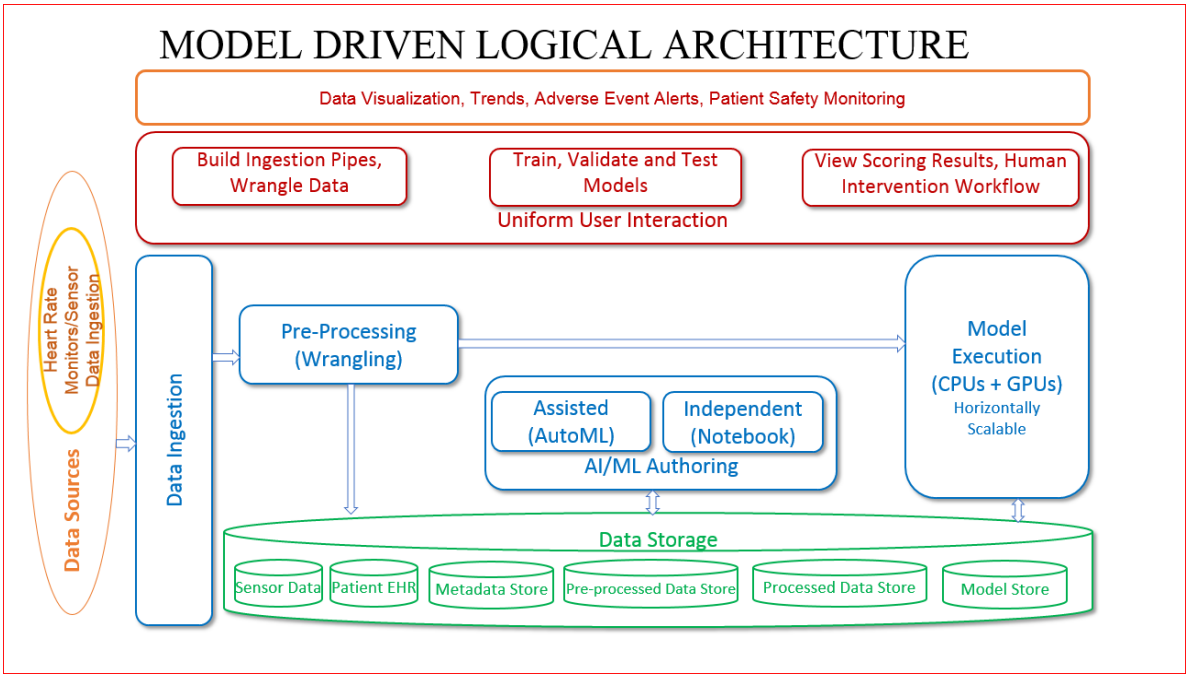


Figure 6. Model Driven Logical Architecture explains data capture and workflow.

c) How can we integrate sensor data in the Clinical Scenario?
This can be explained using the following example (Figure 7: where patient generated data and trial data merge to form the clinical database, to be mapped through CDASH to SDTM/ADAM data.

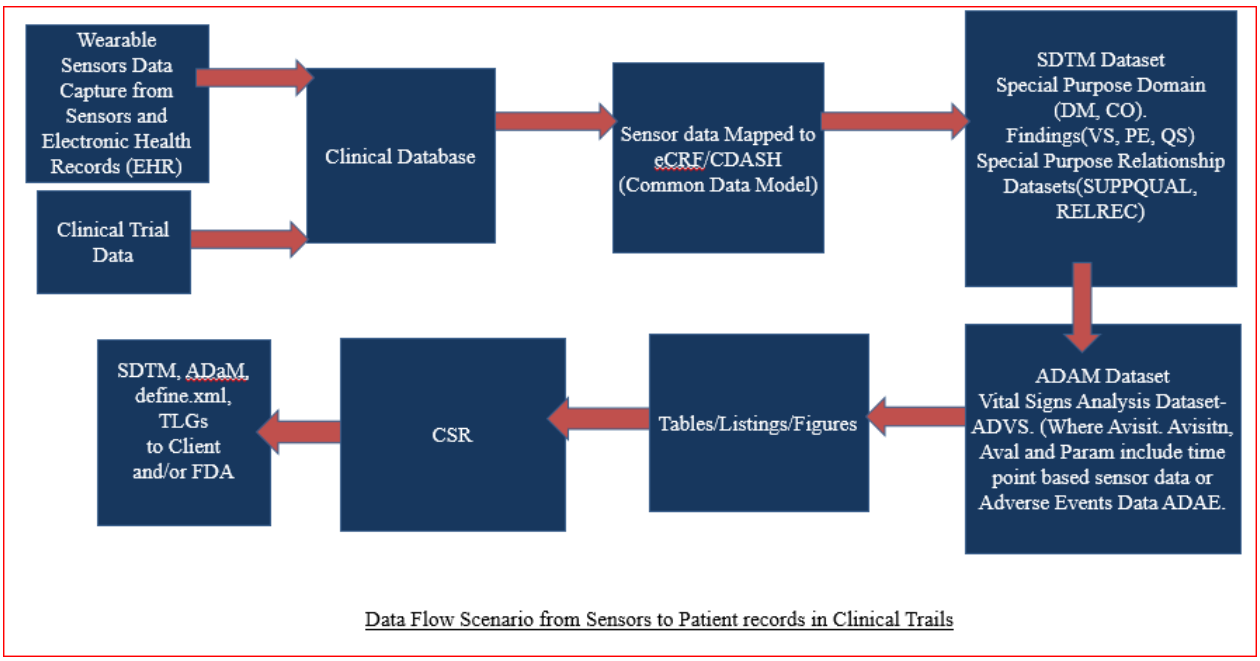


Figure 7. Clinical Scenario Data Workflow: Integrating Sensor and Clinical Trial Data.

d) Reading Sensor data: Sensor data can sourced as:

1. Flat files from Patient Generated wearable devices
2. Sensor data integration with patient Electronic Health Records.

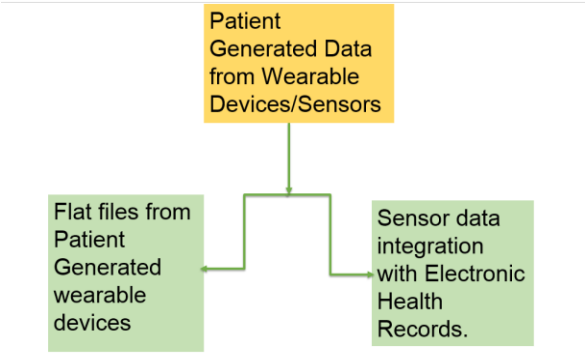


Figure 8. Extracting Patient Generated Data from Wearable Sensors

i. Case One: Flat files from Patient Generated wearable devices:
The following table illustrates how CDASH/CRF Questionnaires that can be associated and mapped with wearable sensor dataset variable. Data can be harmonized to usable format in the clinical context and converted to SDTM/ADAM data structure.

CDASH CRF Label/Question	**Sensor Captured Data**	**Clinical Database Variable Name(CDASH variables Highlighted)**	**Definition**
Date of Vital Signs Measurements	Date	VSDT	Date of measurements.
Sponsor Defined Identifier	ID	VSSPID	Sponsor defined reference number (*For sensor data it could be derived based on user ID*).
Study Day if Vital Signs	Day	VISITDY	Study day of measurements, measured as integer days.
Planned Time Point	Daily Average	VSTPT	Text description of time when measurements should be taken (*For sensor data it could be the average measurement at equal intervals of time*).
Time of Vital Sign Measurements	Time	VSTM (Note: If collected will be derived into VSDTC.)	Time of measurements.
Vital Sign Test Name	*Steps, Distance, Floors, HR, Activity Calories, Blood Glucose.*	VSTEST	Verbatim name of the test or examination used to obtain the measurement or finding (*For sensor data this could be any sensor endpoint that is measured*).
Vital Status	*Missing*	VSSTAT	Used to indicate that a vital signs measurement was not done.
Vital Sign Test Result or Finding	*Result*	VSORRES	Result of the vital signs measurement as originally received or collected (*For sensor data, this could be transposed*

			data, where values are derived in to a single variable for all measurements).
Original Units	Units	VSORRESU	Original units in which the data were collected (This is predefined in sensor data).
Location of Vital Signs Measurement	*Location*	VSLOC	Location on body where measurement was performed (*For sensor data it could be derived as location of sensor for measurement: Example: Write for Fitbit*).
Position of Subject	*Activity*	VSPOS	Position of the subject during a measurement or examination (*For sensor data this is already recorded as Active, or Resting*).

ii. Case Two: Sensor data integration with patient Electronic Health Records.
Standardizing Electronic Health records data to usable format remains a challenge due to the unstructured, disorganized nature of data and uncodified variables. Raw data from Electronic Health Records data is converted to *.xpt files to be converted to CDASH data and SDTM. Some of the main challenges include ascertaining the change in dosage, visit variable and time points including missing values. Variable definitions and derivations from the CRF are used to create CDASH dataset before SDTM dataset is prepared. Some of the CRF, CDASH and SDTM variables from the VS: Highly Recommended Data Collection Fields are listed below:

CRF /Electronic Health records Data Collection Fields	**CDASH Data Collection Field Name**	**SDTM Variable name**
Patient's Weight	VSWGHT	VSORRES, VSTEST
Patient's Systolic Blood Pressure	VSSYSBP	VSORRES, VSTEST
Patient's Diastolic Blood pressure	VSDIABP	VSORRES, VSTEST
Patient's Pulse rate	VSPULSE	VSORRES, VSTEST
Patient's Heart rate	VSHR	VSORRES, VSTEST
Patient's Blood Pressure Position	VSBPPOS	VSPOS
Patient's Weight/Units	VSWGHTU	VSORRESU
Patient's Blood Pressure Units	VSBPU	VSORRESU
Patient's Pulse Units	VSPULSU	VSORRESU
Patient's Heart Rate Units	VSHRU	VSORRESU
Patient's Height	VSHGHT	VSORRES, VSTEST

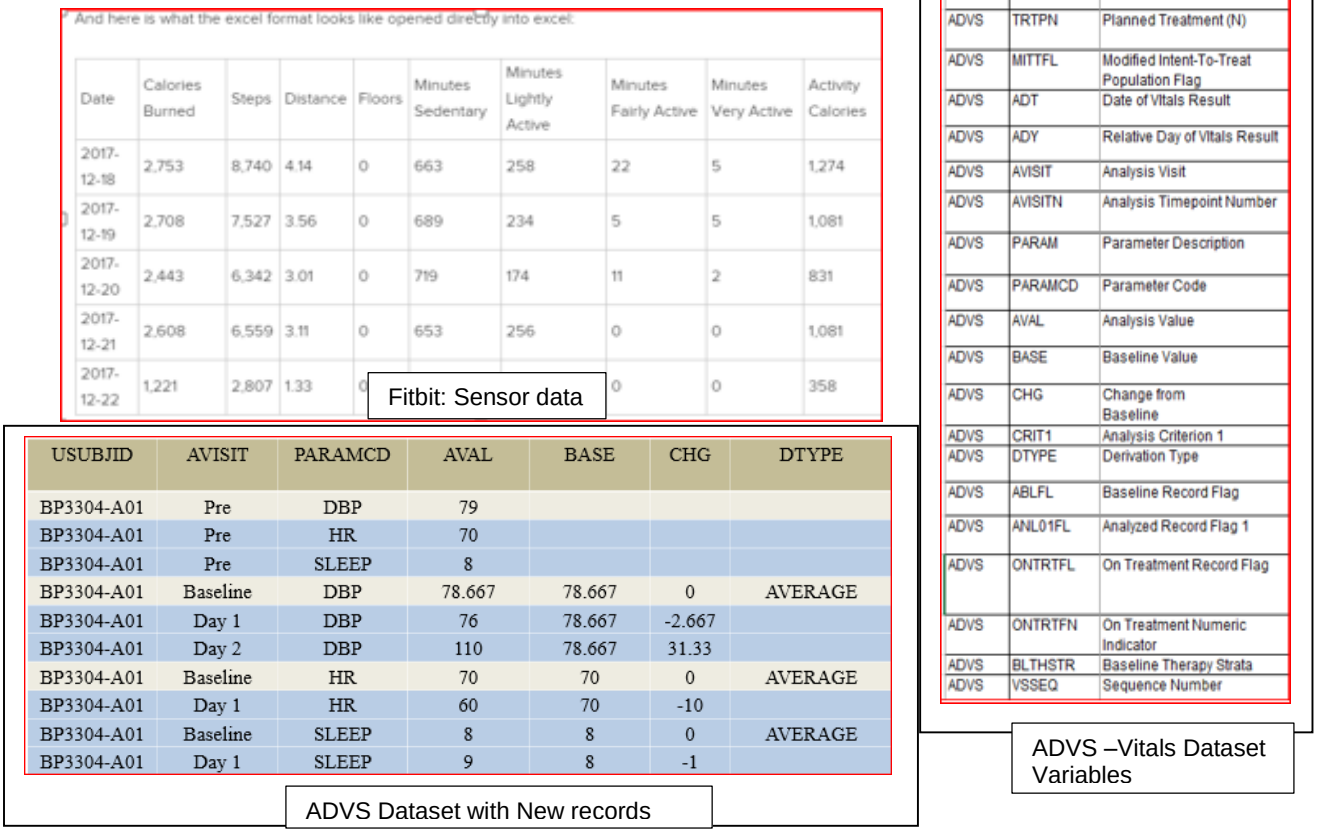
e) Challenges in converting sensor data to standardized clinical data format:

i. Reading unstructured sensor data to patient Electronic Health Records and making it analysis ready needs a lot of data processing. Patient records from Electronic Health Records are converted to flat files and variables are mapped from the CRF to CDASH for enabling conversion to SDTM datasets. The CRF is the only source for variables definition. Some of the main challenges in this method are missing data and locating the unique patient id and visit variables.

ii. Sensor data is semi-structured or unstructured data and it needs to conform to clinical data standards. Below is a snapshot of daily Fitbit data downloaded as an excel spreadsheet and a hypothetical process of merging it with trial data.

Since sensor data is a time series data, data capture can be done in between subject visits by dividing data points between visits, at specific time points (Blue highlighted in the table below). Sensor data in ADVS Dataset:

The variable PARAMCD can include additional coding for sensor captured data. Ex: HR, SBP, DBP, STEPS, SLEEP.



And here is what the excel format looks like opened directly into excel:

Date	Calories Burned	Steps	Distance	Floors	Minutes Sedentary	Minutes Lightly Active	Minutes Fairly Active	Minutes Very Active	Activity Calories
2017-12-18	2,753	8,740	4.14	0	663	258	22	5	1,274
2017-12-19	2,708	7,527	3.56	0	689	234	5	5	1,081
2017-12-20	2,443	6,342	3.01	0	719	174	11	2	831
2017-12-21	2,608	6,559	3.11	0	653	256	0	0	1,081
2017-12-22	1,221	2,807	1.33	0			0	0	358

Fitbit: Sensor data

USUBJID	AVISIT	PARAMCD	AVAL	BASE	CHG	DTYPE
BP3304-A01	Pre	DBP	79			
BP3304-A01	Pre	HR	70			
BP3304-A01	Pre	SLEEP	8			
BP3304-A01	Baseline	DBP	78.667	78.667	0	AVERAGE
BP3304-A01	Day 1	DBP	76	78.667	-2.667	
BP3304-A01	Day 2	DBP	110	78.667	31.33	
BP3304-A01	Baseline	HR	70	70	0	AVERAGE
BP3304-A01	Day 1	HR	60	70	-10	
BP3304-A01	Baseline	SLEEP	8	8	0	AVERAGE
BP3304-A01	Day 1	SLEEP	9	8	-1	

ADVS Dataset with New records

ADVS	STUDYID	Study ID
ADVS	USUBJID	Unique Subject Identifier
ADVS	SITEID	Study Site Identifier
ADVS	TRTP	Planned Treatment
ADVS	TRTPN	Planned Treatment (N)
ADVS	MITTFL	Modified Intent-To-Treat Population Flag
ADVS	ADT	Date of Vitals Result
ADVS	ADY	Relative Day of Vitals Result
ADVS	AVISIT	Analysis Visit
ADVS	AVISITN	Analysis Timepoint Number
ADVS	PARAM	Parameter Description
ADVS	PARAMCD	Parameter Code
ADVS	AVAL	Analysis Value
ADVS	BASE	Baseline Value
ADVS	CHG	Change from Baseline
ADVS	CRIT1	Analysis Criterion 1
ADVS	DTYPE	Derivation Type
ADVS	ABLFL	Baseline Record Flag
ADVS	ANL01FL	Analyzed Record Flag 1
ADVS	ONTRTFL	On Treatment Record Flag
ADVS	ONTRTFN	On Treatment Numeric Indicator
ADVS	BLTHSTR	Baseline Therapy Strata
ADVS	VSSEQ	Sequence Number

ADVS –Vitals Dataset Variables

Figure 9. Example: Mapping Sensor Data to ADVS Dataset and Creating Additional Records.

CONCLUSION

The purpose of this paper is to encourage the including of a 360 degree approach to patient care by adding more data points between visits during a clinical trial. Patient care should widen beyond the date of visit to healthcare provider to include real-time, real world data. Inclusion of Patient reported outcomes and medical devices data would enable appropriate medical care, medication adjustments and interventions at the right time. With the advancement of telemedicine the future promises close monitoring of patient health and shortening the time of drug approval.

REFERENCES

1. HealthITAnalytics. (2018, September 04). FDA Sets Goals for Big Data, Clinical Trials, Artificial Intelligence. Retrieved from https://healthitanalytics.com/news/fda-sets-goals-for-big-data-clinical-trials-artificial-intelligence
2. What information does an electronic health record (EHR) contain? (n.d.). Retrieved from https://www.healthit.gov/faq/what-information-does-electronic-health-record-ehr-contain

3. Health indicator. (2019, February 07). Retrieved from https://en.wikipedia.org/wiki/Health_indicator
4. Health Indicators: A 4-Part Webinar Series. (n.d.). Retrieved from https://www.nlm.nih.gov/nichsr/healthindicators/index.html
5. Explore scientific, technical, and medical research on ScienceDirect. (n.d.). Retrieved from https://www.sciencedirect.com/
6. Lonini, L., Dai, A., Shawen, N., Simuni, T., Poon, C., Shimanovich, L., . . . Jayaraman, A. (2018, November 23). Wearable sensors for Parkinson's disease: Which data are worth collecting for training symptom detection models. Retrieved from https://www.nature.com/articles/s41746-018-0071-z
7. Izmailova, E. S., Wagner, J. A., & Perakslis, E. D. (2018, April 02). Wearable Devices in Clinical Trials: Hype and Hypothesis. Retrieved from https://ascpt.onlinelibrary.wiley.com/doi/10.1002/cpt.966
8. De Arriba-Pérez, F., Caeiro-Rodríguez, M., & Santos-Gago, J. M. (2016, September 21). Collection and Processing of Data from Wrist Wearable Devices in Heterogeneous and Multiple-User Scenarios. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5038811/
9. PCORnet: The National Patient-Centered Clinical Research Network. (2019, May 14). Retrieved from https://www.pcori.org/research-results/pcornet-national-patient-centered-clinical-research-network
10. The Body Map Of Digital Health Sensors - The Medical Futurist. (n.d.). Retrieved from https://medicalfuturist.com/the-body-map-of-digital-health-sensors
11. New digital pills send health data to your doctor – from inside you. (2019, January 23). Retrieved from https://www.scmp.com/lifestyle/health-wellness/article/2183181/new-digital-pills-containing-sensors-send-health-data
12. EMERGING DIGITAL HEALTH TECHNOLOGIES - fda.gov. (n.d.). Retrieved from https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/PatientEngagementAdvisoryCommittee/UCM626406.pdf
13. Sensorlink Archives. (n.d.). Retrieved from https://www.medidata.com/en/product-focus/sensorlink/
14. Exporting / Importing Data from an EHR Using CCDA | DrChrono Blog. (2019, March 08). Retrieved from https://blog.drchrono.com/exporting-importing-data-electronic-health-record-vendor/
15. CDASH. (2017, September 25). Retrieved from https://www.cdisc.org/standards/foundational/cdash
16. CDISC. (n.d.). Retrieved from https://www.cdisc.org/

ACKNOWLEDGMENTS

Thank you EG Life Sciences for providing the opportunity to work on this paper. Special mention to Poornima Gurram, EG Life Sciences for reviewing this paper and for sharing her experience with patient Electronic Health Records.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Surabhi Dutta
EG Life Sciences
400 Atrium Drive
Somerset, NJ-08873