Paper PP18

Incorporating Digital Health Technology (DHT) Data in SDTM – Challenges and Examples

Sandra VanPelt Nguyen, Pfizer Inc., USA
Swarupa Sudini, Pfizer Inc., USA

ABSTRACT

As digital health technologies (DHT) like wearables are increasingly used to capture data in clinical trials, sponsors face challenges in integrating and mapping this data to CDISC data standards for regulatory submissions. The diversity of data sources, high granularity, large volumes, use of software algorithms, and inclusion of supportive contextual data complicate this process. Current CDISC guidance offers limited examples, making standardization difficult. Some data may not fit clearly into existing SDTM domains, while other data presents challenges related to how it's collected or reported by devices or vendors. This paper will examine some of the challenges in mapping DHT data, summarize the process used at Pfizer to determine the target mapping, and share examples illustrating these challenges.

INTRODUCTION

Digital Health Technology, as defined by the U.S. Food and Drug Administration[1], is a system that uses computing platforms, connectivity, software, and/or sensors, for health care and related uses. This can include hardware and/or software, may qualify as a device under federal law[2], and may involve active or passive data collection. Some examples of DHT used at Pfizer include mobile apps for collection of symptoms, voice assessments, and events such as migraine episodes which may also include data regarding platform, operating system, and other app details. Various activity trackers have been used to collect sleep and activity metrics for which we've also collected data on device usage, algorithms, and data loss. We've tested a wearable patch to collect ECG and vital signs data, along with data on activity and posture, and an insole device for gait measurements as well as posture, spatial, temporal, and pressure data. We'll be sharing examples for some of these devices.

Examples of DHT at Pfizer:

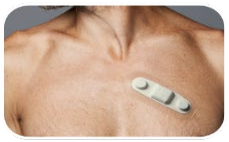
Mobile App

- Daily symptoms
- Daily voice assessments
- Enrollment and end of study questionnnaires
- Platform, OS, app details

Activity Trackers

- Sleep & activity metrics (raw & derived)
- Device usage metrics
- Algorithm versions
- Data loss

Wearable Patch



- ECG
- Vital signs
- Activity tracking
- Posture

Insole Device

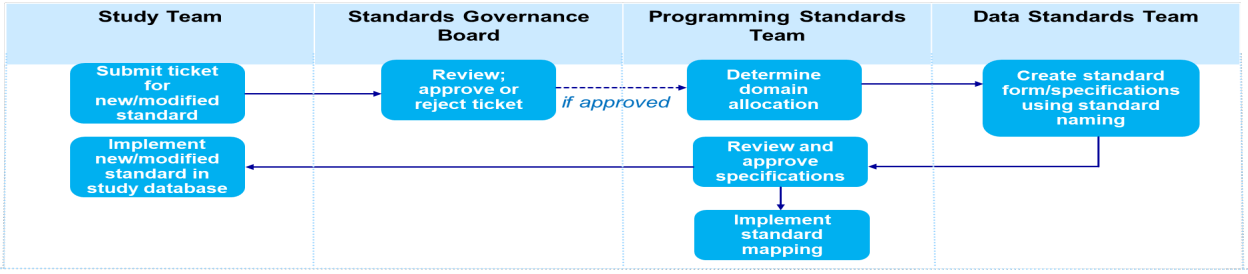
- Gait parameters (raw, derived, estimated)
- Posture
- Spatial, temporal, and pressure data

Studies may use or assess more than one DHT device. This may be done to compare the effectiveness, accuracy, and reliability of various devices or to qualify devices for future studies. In these types of studies, there is extra attention on device and usage issues such as battery life, device syncing issues, amount of data loss, and comfort and wear time metrics. It's critical that participants can wear and use the devices properly so that we can get accurate and complete data.

DOMAIN ALLOCATION PROCESS

DHT data is generally aggregated by a vendor and transferred to Pfizer using our non-CRF data process, similar to data from a central lab. Our Programming Standards team assesses the data being collected and identifies the appropriate target SDTM domain(s) which then drives end-to-end standards design and facilitates automation. The "domain allocation" process often requires collaboration with the study's Clinical and Statistical team members to get an understanding of the clinical significance and interpretation of the data, how various data points in the study relate to each other, and the plans for analysis. For new data types not covered in the SDTM Implementation Guides (SDTMIGs)[3], we'll consult other resources such as the Medical Device Implementation Guide[4], Therapeutic Area User Guides (TAUGs)[5], or consider draft domains/terminology or potentially custom domains.

Once the target SDTM domain(s) are identified, Programming Standards works with the Data Standards team to develop CDASH-based data file specifications for the vendor to utilize when providing the non-CRF data. The implementation of standard naming conventions for data collection fields allows us to standardize and automate SDTM extraction and mapping programs with minimal manipulation of the data.



CHALLENGES AND EXAMPLES

DHT data collection usually involves a device, or potentially a combination of devices, and data collected by the device(s) or about the device(s) often gets split out to multiple SDTM domains but is highly related and may need to be brought together for analysis. Additional operational data not typically seen with CRF-collected data is often included and must be assessed for relevance and context; anything not needed for analysis and reporting or as supportive information for interpreting the data may not need to be mapped to SDTM. Some of the data may not have an obvious standard domain, thus requiring research, cross-functional collaboration, and sometimes a bit of creativity to determine an appropriate mapping. Furthermore, many wearables will continuously collect data, which can lead to large volumes of data to transfer and analyze. Organizations must determine the appropriate level of summarization to include in SDTM.

Common challenges include:

1) DEVICE DATA AND ITS ASSOCIATION TO PARTICIPANTS/RESULTS

As most DHT data is collected via a device (or combination of devices), Medical Device domains should be created; minimally the DI (Device Identifiers) domain. Organizations must first determine what qualifies as a device or "Software as a Medical Device" (SaMD) according to Section 201(h)(1) of the U.S. Food, Drug, and Cosmetic (FD&C) Act[2] and assign identifiers accordingly. Identifier variable SPDEVID (Sponsor Device Identifier) is critical for linking data from the same device across domains. The DI domain must contain the "Type" parameter (DIPARMCD = "TYPE") at a minimum, but can additionally include information such as the manufacturer, model, trade name, and serial number, depending on the level of detail needed to understand the data and differentiate between devices. Other information regarding properties and settings of the devices may need to be mapped to the DO (Device Properties) or DU (Device In-Use) domain, with information on device usage in DX (Device Exposure) and any device-related events (such as malfunctions) in the DE (Device Events) domain if collected.

Additionally, the RELDEV (Relating Devices) domain (draft for SDTMIG v4.0) may be used to show how different devices relate to each other (e.g. parent/child relationships when there are device components which are part of a larger device/system).

Example 1: a study took measurements using two wrist-worn accelerometers (left and right wrists) on the participant. Results were not associated with an individual accelerometer but were derived from the combination of left and right wrist data. Each accelerometer is identified as a device, with the combination of the two being a device "system" which was used for SPDEVID across other domains that contain the device-derived data.

While data in the MK (Musculoskeletal System Findings) domain here attributes results to the sensor "system", RELDEV shows the parent-child relationship between the sensor system (LEVEL=1 for parent) and the individual left and right sensors (LEVEL=2 for child components). The DI domain includes records for the parent and its components.

di.xpt

STUDYID	DOMAIN	SPDEVID	DIPARMCD	DIPARM	DIVAL
ABC-123	DI	Scratch Sensor System	DEVTYPE	Device Type	Kinesiology ambulatory recorder system
ABC-123	DI	Right Sensor	DEVTYPE	Device Type	Kinesiology ambulatory recorder
ABC-123	DI	Right Sensor	WEARLOC	Wear Location	Right Wrist
ABC-123	DI	Left Sensor	DEVTYPE	Device Type	Kinesiology ambulatory recorder
ABC-123	DI	Left Sensor	WEARLOC	Wear Location	Left Wrist

reldev.xpt

STUDYID	SPDEVID	PARENT	LEVEL
ABC-123	Scratch Sensor System		1
ABC-123	Right Sensor	Scratch Sensor System	2
ABC-123	Left Sensor	Scratch Sensor System	2

mk.xpt

STUDYID	DOMAIN	USUBJID	SPDEVID	MKORRES	MKORRESU
ABC-123	MK	1001	Scratch Sensor System	220	
ABC-123	MK	1001	Scratch Sensor System	14.3	MINUTES

Example 2: During a study that used an activity tracking device, a malfunction occurred due to a firmware bug, resulting in data loss for some participants. The study wanted to capture the amount of data lost to incorporate in the analysis. The Device Events (DE) domain was used to capture these malfunctions of the device, with the associated amount of data loss in FADE (Findings About Device Events).

de.xpt

STUDYID	DOMAIN	USUBJID	SPDEVID	DETERM	DEDECOD	DECAT	VISIT	DESTDTC
XYZ123	DE	1001	1234	Device Data Loss	Device Data Issue	MALFUNCTION	VISIT1	2020-12-04T14:14:19

fade.xpt

STUDYID	DOMAIN	USUBJID	SPDEVID	FATEST	FAOBJ	FACAT	FASCAT	FAORRES	FAORRESU	VISIT	FADTC
XYZ123	FA	1001	1234	Amount of Data loss	Device Data Issue	PHYSICAL ACTIVITY MONITORING	SLOW WALK	0.871163163	fraction of 1	VISIT1	2020-12-04T14:14:19

2) IDENTIFYING APPROPRIATE SDTM DOMAINS

As we work through our domain allocation process for new data standards and try to identify the target SDTM domains, we often find that there are not clear standard domains documented within current guidance. A Findings About (FA) domain or custom domain may ultimately be used when nothing else seems to fit.

Example 3: a study collected voice recordings via a mobile app which were used to assess pitch, power, and spectral structure of speech. There were no examples related to vocal/speech assessments or any related controlled terminology to help us identify a domain. At the time, limited body system-based domains were published, and the data did not seem to fit into any of them based on the available examples and existing terminology, so a custom Findings domain (ZV) was created. Even now, discussion is underway within the CDISC DHT team to determine the most appropriate mapping of such data, as multiple body systems are involved with speech.

zv.xpt

STUDYID	USUBJID	ZVGRPID	ZVTEST	ZVCAT	ZVORRES	ZVORRESU	ZVDTC	ZVENDTC
ABC123	1001	AHH	Max Phonation Time	VOICE ACOUSTICS - SUMMARY DATA	8.308	sec	2021-04-20T11:46:08	2021-04-20T11:47:12
ABC123	1001	EE	Cepstral Peak Prominence	VOICE ACOUSTICS - SUMMARY DATA	6.30802	dB	2021-04-20T11:46:55	2021-04-20T11:48:17
ABC123	1001	EE	Coefficient of Variation F0	VOICE ACOUSTICS - SUMMARY DATA	0.02692		2021-04-20T11:46:55	2021-04-20T11:48:17

Example 4: a low-interventional study designed to evaluate the performance of various gait and activity monitoring technologies in healthy subjects with one of the devices under study being connected insoles with built-in pressure and movement sensors. An assessment across various CDISC guidance documents identified a related example in the Type 1 Diabetes - Exercise and Nutrition Modules[6] TAUG, section 4.3 “Wearable Devices Measuring Exercise Parameters”, which utilizes PR (Procedures), FA, (Findings About) and DI domains for information about the activity monitoring, findings about the intervention exercise “not related to a specific body system domain”, and the wearable device respectively.

e-Data

STUDY	SUBJECT	VISIT	CATEGORY	PROCEDURE	SPDEVID	DEVTYPE	DATE OF COLL	LOC	ACMTEST	RESULT	UNIT
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	SLOW WALK	1234	FEETME-INSOLES	2020-12-04T14:14:19	FOOT	Single Limb Support Duration	0.190867475	sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	SLOW WALK	1234	FEETME-INSOLES	2020-12-04T14:14:19	FOOT	Double Support Duration	0.102842699	sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	SLOW WALK	1234	FEETME-INSOLES	2020-12-04T14:14:19	FOOT	Swing Duration	0.667083016	sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	SLOW WALK	1234	FEETME-INSOLES	2020-12-04T14:14:19	FOOT	Stance Duration	0.747797779	sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	SLOW WALK	1234	FEETME-INSOLES	2020-12-04T14:14:19	FOOT	Stride Length	0.04	m
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	SLOW WALK	1234	FEETME-INSOLES	2020-12-04T14:14:19	FOOT	Step Duration	0.504635666	sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	SLOW WALK	1234	FEETME-INSOLES	2020-12-04T14:14:19	FOOT	Stride Duration	0.780439629	sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	SLOW WALK	1234	FEETME-INSOLES	2020-12-04T14:14:19	FOOT	Average Cadence	0.070867475	/min
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	SLOW WALK	1234	FEETME-INSOLES	2020-12-04T14:14:19	FOOT	Gait Speed	0.02	m/sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	SLOW WALK	1234	FEETME-INSOLES	2020-12-14T14:14:19	FOOT	Data Loss	0.871163163	fraction of 1
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	OUTSIDE LAB	1234	FEETME-INSOLES	2020-12-04T15:08:59	FOOT	Single Limb Support Duration	40.16513928	sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	OUTSIDE LAB	1234	FEETME-INSOLES	2020-12-04T15:08:59	FOOT	Double Support Duration	21.91953968	sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	OUTSIDE LAB	1234	FEETME-INSOLES	2020-12-04T15:08:59	FOOT	Swing Duration	40.20087319	sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	OUTSIDE LAB	1234	FEETME-INSOLES	2020-12-04T15:08:59	FOOT	Stance Duration	62.14287681	sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	OUTSIDE LAB	1234	FEETME-INSOLES	2020-12-04T15:08:59	FOOT	Stride Length	1.167348924	m
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	OUTSIDE LAB	1234	FEETME-INSOLES	2020-12-04T15:08:59	FOOT	Step Duration	0.5109375	sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	OUTSIDE LAB	1234	FEETME-INSOLES	2020-12-04T15:08:59	FOOT	Stride Duration	1.0234375	sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	OUTSIDE LAB	1234	FEETME-INSOLES	2020-12-04T15:08:59	FOOT	Average Cadence	58.6259542	/min
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	OUTSIDE LAB	1234	FEETME-INSOLES	2020-12-04T15:08:59	FOOT	Gait Speed	1.140680587	m/sec
XYZ123	1001	VISIT1	PHYSICAL ACTIVITY MONITORING	OUTSIDE LAB	1234	FEETME-INSOLES	2020-12-04T15:08:59	FOOT	Data Loss	0.253350555	fraction of 1

pr.xpt

STUDYID	DOMAIN	USUBJID	SPDEVID	PRTRT	PRCAT	PRLOC	VISIT	PRSTDTC
XYZ123	PR	1001	1234	SLOW WALK	PHYSICAL ACTIVITY MONITORING	FOOT	VISIT1	2020-12-04T14:14:19
XYZ123	PR	1001	1234	OUTSIDE LAB	PHYSICAL ACTIVITY MONITORING	FOOT	VISIT1	2020-12-04T15:08:59

fapr.xpt

STUDYID	DOMAIN	USUBJID	SPDEVID	FATEST	FAOBJ	FACAT	FAORRES	FAORRESU	VISIT	FADTC
XYZ123	FA	1001	1234	Single Limb Support Duration	SLOW WALK	PHYSICAL ACTIVITY MONITORING	0.190867475	sec	VISIT1	2020-12-04T14:14:19
XYZ123	FA	1001	1234	Double Support Duration	SLOW WALK	PHYSICAL ACTIVITY MONITORING	0.102842699	sec	VISIT1	2020-12-04T14:14:19
XYZ123	FA	1001	1234	Swing Duration	SLOW WALK	PHYSICAL ACTIVITY MONITORING	0.667083016	sec	VISIT1	2020-12-04T14:14:19
XYZ123	FA	1001	1234	Stance Duration	SLOW WALK	PHYSICAL ACTIVITY MONITORING	0.747797779	sec	VISIT1	2020-12-04T14:14:19
XYZ123	FA	1001	1234	Stride Length	SLOW WALK	PHYSICAL ACTIVITY MONITORING	0.04	m	VISIT1	2020-12-04T14:14:19
XYZ123	FA	1001	1234	Step Duration	SLOW WALK	PHYSICAL ACTIVITY MONITORING	0.504635666	sec	VISIT1	2020-12-04T14:14:19
XYZ123	FA	1001	1234	Stride Duration	SLOW WALK	PHYSICAL ACTIVITY MONITORING	0.780439629	sec	VISIT1	2020-12-04T14:14:19
XYZ123	FA	1001	1234	Average Cadence	SLOW WALK	PHYSICAL ACTIVITY MONITORING	0.070867475	/min	VISIT1	2020-12-04T14:14:19
XYZ123	FA	1001	1234	Gait Speed	SLOW WALK	PHYSICAL ACTIVITY MONITORING	0.02	m/sec	VISIT1	2020-12-04T14:14:19

di.xpt

STUDYID	DOMAIN	SPDEVID	DISEQ	DIPARMCD	DIPARM	DIVAL
XYZ123	DI	1234	1	DEVTYPE	Device Type	INSOLES
XYZ123	DI	1234	2	TRADENAM	Trade Name	FEETME

3) LINKING NON-CRF DHT DATA WITH CRF DATA

Differences in data structure and collection timing across domains (for example, continuous device data versus CRF data assessed at a specified point in time) can complicate record-level linkage. If the non-CRF data is related to CRF

data that we're collecting, we need to ensure we can link that data together for downstream review and analysis, which can be challenging if we don't have common identifiers.

Example 5: This study utilized a mobile app and wearable devices to detect changes in vital signs across different phases of the menstrual cycle, identify and characterize migraine episodes, and track temperatures to help predict and manage menstrual-related migraine attacks. Menstrual cycle start and end dates were recorded in the CRF, with skin temperature continuously collected using a wearable device, and participants entered information pertaining to migraine episodes in the app. Menstrual cycle data collected on the CRF was mapped to the RP (Reproductive System Findings) domain, skin temperature to the VS (Vital Signs) domain, and migraine episodes to the CE (Clinical Events) domain.

Due to the structural and timing differences between the CRF and non-CRF data, direct linking within SDTM was not feasible. Therefore, the recommendation was to use windowing in ADaM to establish connections between menstrual cycle phases and the corresponding temperatures and migraine episodes.

Menstrual Cycle CRF data

Cycle	Start Date	End Date
1	2024-01-01	2024-01-04
2	2024-02-01	2024-02-05
3	2024-03-01	2024-03-04

rp.xpt

STUDYID	DOMAIN	USUBJID	RPSPID	RPGRPID	RPTESTCD	RPTEST	RPORRES
ABC234	RP	1001	1	1	MCYSTDTC	Start Date of Menstrual Cycle	2024-01-01
ABC234	RP	1001	2	1	MCYENDTC	End Date of Menstrual Cycle	2024-01-04
ABC234	RP	1001	3	2	MCYSTDTC	Start Date of Menstrual Cycle	2024-02-01
ABC234	RP	1001	4	2	MCYENDTC	End Date of Menstrual Cycle	2024-02-05
ABC234	RP	1001	5	3	MCYSTDTC	Start Date of Menstrual Cycle	2024-03-01
ABC234	RP	1001	6	3	MCYENDTC	End Date of Menstrual Cycle	2024-03-04

e-Data

STUDY	SUBJECT	DATE OF COLL	SPDEVID	LOCATION	CATEGORY	TEST	RESULT	UNIT	EVAL	DEVTYPE	TRADENAM	MANUF
ABC234	1001	2024-01-01T09:30:00	1111	WRIST JOINT	VITAL SIGNS MONITORING	Skin Temperature	35.0598023	C		WEARABLE	EmbracePlus	EMPATICA
ABC234	1001	2024-01-01T09:31:00	1111	WRIST JOINT	VITAL SIGNS MONITORING	Skin Temperature	34.87848856	C		WEARABLE	EmbracePlus	EMPATICA
ABC234	1001	2024-01-01T09:32:00	1111	WRIST JOINT	VITAL SIGNS MONITORING	Skin Temperature	34.04824848	C		WEARABLE	EmbracePlus	EMPATICA
ABC234	1001	2024-01-01T09:33:00	1111	WRIST JOINT	VITAL SIGNS MONITORING	Skin Temperature	35.44244938	C		WEARABLE	EmbracePlus	EMPATICA
ABC234	1001	2024-01-01T09:34:00	1111	WRIST JOINT	VITAL SIGNS MONITORING	Skin Temperature	32.24911394	C		WEARABLE	EmbracePlus	EMPATICA
....	....	....	....	....	....	....	....	....		....	....	....
....	....	....	....	....	....	....	....	....		....	....	....
ABC234	1001	2024-01-03T09:38:00	1111	WRIST JOINT	VITAL SIGNS MONITORING	Skin Temperature	34.71985395	C		WEARABLE	EmbracePlus	EMPATICA
ABC234	1001	2024-01-03T09:39:00	1111	WRIST JOINT	VITAL SIGNS MONITORING	Skin Temperature	33.91677987	C		WEARABLE	EmbracePlus	EMPATICA
ABC234	1001	2024-01-03T09:40:00	1111	WRIST JOINT	VITAL SIGNS MONITORING	Skin Temperature	34.51597354	C		WEARABLE	EmbracePlus	EMPATICA
ABC234	1001	2024-01-03T09:41:00	1111	WRIST JOINT	VITAL SIGNS MONITORING	Skin Temperature	35.45149701	C		WEARABLE	EmbracePlus	EMPATICA
....	....	....	....	....	....	....	....	....		....	....	....
....	....	....	....	....	....	....	....	....		....	....	....
ABC234	1001	2024-01-03T00:00:00	2222		MIGRAINE MONITORING	Moderate Migraine Episode Date	2024-01-03		SUBJECT	APP	MIRA	

vs.xpt

STUDYID	DOMAIN	USUBJID	SPDEVID	VSTESTCD	VSTEST	VSCAT	VSORRES	VSORRESU	VSLOC	VSDTC
ABC234	VS	1001	1111	TEMP	Temperature	VITAL SIGNS MONITORING	35.059802	C	WRIST JOINT	2024-01-01T09:30:00
ABC234	VS	1001	1111	TEMP	Temperature	VITAL SIGNS MONITORING	34.878489	C	WRIST JOINT	2024-01-01T09:31:00
ABC234	VS	1001	1111	TEMP	Temperature	VITAL SIGNS MONITORING	34.048248	C	WRIST JOINT	2024-01-01T09:32:00
ABC234	VS	1001	1111	TEMP	Temperature	VITAL SIGNS MONITORING	35.442449	C	WRIST JOINT	2024-01-01T09:33:00
ABC234	VS	1001	1111	TEMP	Temperature	VITAL SIGNS MONITORING	32.249114	C	WRIST JOINT	2024-01-01T09:34:00
....	....	....	....	....	....	....	....	....	....	....
....	....	....	....	....	....	....	....	....	....	....
ABC234	VS	1001	1111	TEMP	Temperature	VITAL SIGNS MONITORING	34.719854	C	WRIST JOINT	2024-01-03T09:38:00
ABC234	VS	1001	1111	TEMP	Temperature	VITAL SIGNS MONITORING	33.91678	C	WRIST JOINT	2024-01-03T09:39:00
ABC234	VS	1001	1111	TEMP	Temperature	VITAL SIGNS MONITORING	34.515974	C	WRIST JOINT	2024-01-03T09:40:00
ABC234	VS	1001	1111	TEMP	Temperature	VITAL SIGNS MONITORING	35.451497	C	WRIST JOINT	2024-01-03T09:41:00

ce.xpt

STUDYID	DOMAIN	USUBJID	SPDEVID	CETERM	CESEV	CESTDTC
ABC234	CE	1001	2222	Migraine	MODERATE	2024-01-03

di.xpt

STUDYID	DOMAIN	SPDEVID	DISEQ	DIPARMCD	DIPARM	DIVAL
ABC234	DI	1111	1	DEVTYPE	Device Type	Wearable
ABC234	DI	1111	2	MANUF	Manufacturer	EMPATICA
ABC234	DI	1111	2	TRADENAM	Trade Name	EmbracePlus
ABC234	DI	2222	1	DEVTYPE	Device Type	APP
ABC234	DI	2222	2	TRADENAM	Trade Name	MIRA

4) METRICS AND OTHER OPERATIONAL DATA

Data related to operational and technical aspects of the device and its usage can be challenging to map to standard domains and may not be required for analysis and reporting. Cross-functional collaboration is often necessary to learn the intended use of the data, determine whether it is needed to interpret the data appropriately, and establish how it relates to other collected data.

Example 6: a study that used a wearable device to assess physical activity and performance outcomes discovered unexpected occurrences of data loss that were due to factors such as battery issues, insufficient wear time, and memory usage which were not necessarily a malfunction of the device itself. To proactively monitor these usage issues, the team implemented real-time monitoring of the device and the data flow from the device to a mobile app. The data was then transmitted to a server, with a dashboard programmed to summarize various usage metrics. This enabled the quantification of data loss based on the dashboard parameters, providing a clear understanding of where and why data gaps were occurring so that sites and participants could be notified and educated on how to properly use the devices.

We initially explored the use of DU and DO domains, however the data did not align well with the definitions of either domain [*DU is intended to represent properties of the study device that are deliberately configured within the context of the study, while DO captures key characteristics of the device used*]. As a result, we chose to map the dashboard monitoring metrics to a custom Findings domain - Device Status Monitoring (ZD).

e-Data

Study	Subject	Category	Device Name	Status	Device B	Memory	Last Wor	Data Uploaded	Serial	Upload Info Date
ABC123	1001	Device Status Activity	Actigraph	Issue	64%	36%	Not Worn	93%	STM2E24247751	4/8/2025 5:01:57

zd.xpt

STUDYID	DOMAIN	USUBJID	SPDEVID	ZDTESTCD	ZDTEST	ZDORRES	ZDORRESU	ZDDTC
ABC123	ZD	1001	STM2E24247751	STATUS	Status	Issue		2025-04-08T05:01:57
ABC123	ZD	1001	STM2E24247751	DEVBUSG	Device Battery Usage	64	%	2025-04-08T05:01:57
ABC123	ZD	1001	STM2E24247751	MEMUSG	Memory Usage	36	%	2025-04-08T05:01:57
ABC123	ZD	1001	STM2E24247751	DUPLOAD	Data Uploaded	100	%	2025-04-08T05:01:57
ABC123	ZD	1001	STM2E24247751	LASTWORN	Last Worn State	Not Worn		2025-04-08T05:01:57

di.xpt

STUDYID	DOMAIN	SPDEVID	DISEQ	DIPARMCD	DIPARM	DIVAL
ABC123	DI	STM2E24247751	1	DEVTYPE	Device Type	Wearable
ABC123	DI	STM2E24247751	2	MANUF	Manufacturer	Actigraph
ABC123	DI	STM2E24247751	3	SERIAL	Serial Number	STM2E24247751

5) LEVEL OF GRANULARITY

The level of granularity of the data needs to be considered - whether to attain and map all raw source data acquired by the device, for which the volume could be huge, or only bring in and map data which has been aggregated/summarized to some extent. These devices can collect large volumes of data, but you may only be interested in averages over certain time intervals, for example, rather than minute-by-minute or second-by-second. Any aggregation likely involves an algorithm which should be accounted for in the data submission.

Example 7: An atopic dermatitis study assessed frequency and duration of scratching during the night. While data was collected on a continuous basis during the amount of time the device was worn each day, we only needed the quantity and duration of scratch bouts that occurred during each day's major rest period. Algorithms were applied to determine the interval of time associated with the major rest period (the longest period during the 24-hour day in which sleep was the intended activity) and then to identify occurrence of scratch bouts based on arm movements. The CDISC DHT team mapped the data to the MK (Musculoskeletal System Findings) domain based on scratching involving movement. To denote the summarization algorithms applied, the --ANMETH (Analysis Method) variable is used.

mk.xpt:

STUDYID	DOMAIN	USUBJID	SPDEVID	MKTESTCD	MKTEST	MKORRES	MKORRESU	MKLOC	MKLAT	MKMETHOD	MKANMETH	MKDTC	MKENDTC	MKEVINTX
ABC-123	MK	1001	Scratch Sensor System	NSCRATCH	Number of Scratch Bouts	220		WRIST	BOTH	ACTIGRAPHY	PFIZER SCRATCH ALGORITHM	2018-12-01T12:00	2018-12-02T11:59	SLEEP PERIOD
ABC-123	MK	1001	Scratch Sensor System	DSCRATCH	Duration of Scratch Bouts	14.3	MINUTES	WRIST	BOTH	ACTIGRAPHY	PFIZER SCRATCH ALGORITHM	2018-12-01T12:00	2018-12-02T11:59	SLEEP PERIOD

CONCLUSION

As shown, mapping DHT data to CDISC standards is feasible, but it can present several challenges. While some guidance is available to support this work, additional DHT-specific direction and examples would be valuable. The CDISC DHT team has partnered with the Digital Medicine Society (DiMe[7]) and is actively developing and documenting key concepts and best practices related to DHTs, including multiple examples based on digital endpoints from the DiMe library[8]. Several endpoints have already been modeled in SDTM, including device domains, and the team is working to expand the variety of use cases and add corresponding analysis data examples. A DHT portal is also being developed on the CDISC website to house this information, with the goal of addressing common questions and improving standardization across the industry.

REFERENCES

1) U.S. Food and Drug Administration: Digital Health Technologies for Remote Data Acquisition in Clinical Investigations Guidance for Industry, Investigators, and Other Stakeholders (https://www.fda.gov/media/155022/download)

2) U.S. Government: Food, Drug, and Cosmetic (FD&C) Act (https://www.govinfo.gov/content/pkg/COMPS-973/pdf/COMPS-973.pdf)

3) CDISC: SDTMIG (https://www.cdisc.org/standards/foundational/sdtmig)

4) CDISC: SDTMIG for Medical Devices v1.1 (https://www.cdisc.org/standards/foundational/medical-devices-sdtmig/sdtmig-medical-devices-v1-1)

5) CDISC: Therapeutic Area User Guides (https://www.cdisc.org/standards/therapeutic-areas)

6) CDISC: Type 1 Diabetes – Exercise and Nutrition Modules v1.0 (https://www.cdisc.org/system/files/members/standard/ta/TAUG_T1D_Exercise_and_Nutrition_Modules_v1.0.pdf)

7) Digital Medicine Society (DIME): https://dimesociety.org/

8) DiME: Library of Digital Endpoints (https://dimesociety.org/library-of-digital-endpoints/)

9) CDISC: https://www.cdisc.org/

CONTACT INFORMATION

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

Author Name: Sandra VanPelt Nguyen
Company: Pfizer, Inc.
Email: sandra.vanpeltnguyen@pfizer.com

Author Name: Swarupa Sudini
Company: Pfizer, Inc.
Email: swarupa.sudini@pfizer.com