Data Protection Impact Assessment (DPIA)

Date	Version	Author	Reason for amendment
1st June 2024	1.0	Oscar Boldt-Christmas	Create formalised Data Protection Impact Assessment
10th January 2025	2.0	Claire Robinson	Amend Data Protection Impact Assessment to follow NHS guidance.
24th January 2025	2.1	Oliver Åstrand	Improvements to risk assessment

Background

A data protection impact assessment (DPIA) will help you to identify and mitigate potential data protection risks to an acceptable level before using or sharing (processing) data that identifies individuals (personal data).

A DPIA will also help you meet a number of data protection legal requirements including:

- Data protection by design - privacy and data protection issues must be considered at the start, or in the design phase, of a new system, product or process, then continuously while it exists.
- Accountability - your organisation is responsible for showing how it complies with data protection laws.
- Transparency - personal data must be used and shared in a transparent way.
- Security - adequate measures need to be in place to protect data. This can range from policies and procedures to technical security measures such as encryption of data.

DPIAs are mandatory when there is a high risk to individuals, such as when using the health and care data of a large number of people. However, health and care organisations are strongly advised to complete a DPIA when using and sharing personal data in a new or substantially changed way.

A DPIA involves a risk assessment. If a high-level risk remains after applying mitigations, then you must consult with the Information Commissioner’s Office (ICO) for further advice before starting to collect, use or share the data.

A DPIA is a live document - you must update it if there are any changes to:

- the purpose - why you are proposing to use or share personal data
- the manner - how you will use or share the data
- who is involved - the organisations using and sharing personal data

Contents

1 Screening questions

1.1 Do you need to do a DPIA?

The project involves the use of a digital medium (hereinafter referred to as "Support") to create draft medical records from listening to patients' conversations. As such it has been decided a DPIA is required for the following reasons:

- Tandem health will be/is using and sharing data (including special category data).
- Tandem Health are implementing a new technology.
- Risk assessments are necessary to process data.

1.2 Summary of how data will be used and shared

The data collected is primarily derived from patient conversations. The data undergoes transcription, is processed to create draft medical notes, and is then transferred to the electronic medical records system with user approval. Key steps in data handling include:

1. Conversations are segmented into encrypted audio chunks for transcription into text. This audio data is deleted shortly after processing (within 5 minutes).

2. Transcripts are processed using a language model to generate draft notes. These notes are reviewed and edited by the user before transfer.

3. Data and notes are stored temporarily to facilitate troubleshooting and support. After a period, the data is minimised, removing identifiers like names, addresses, and social security numbers.

4. Aggregated, anonymised data may be used for improving the service but is stripped of sensitive identifiers. The storage and processing occur securely on Microsoft Azure servers within the EU.

1.3 Description of the data

☒	Personal data [individuals can be identified]
☒	Pseudonymised data [identifiers, for example name or NHS number, are replaced with a unique number or code (a pseudonym)]
☒	Anonymous data [not identifiable, for example trends or statistics]

Pseudonymisation overview:

Data Minimisation: After 30 days transcripts and notes are subjected to a data minimisation algorithm, ensuring that sensitive identifiers such as names, social security numbers, and addresses are removed. This applies even if such identifiers are mentioned during the call.

Storage and Retention: Minimized data is retained for the duration of the service contract with healthcare professionals or institutions, as well as all data, including logs and backups, being deleted within 30 days of the service contract's termination.

Usage of Pseudonymised Data: Pseudonymised data may be used to generate usage statistics and for improving functionality, and to facilitate troubleshooting regarding system issues.

2 Why do you need the data?

2.1 What are the purposes for using or sharing the data?

The purpose is to utilise a digital medium to create draft medical records from listening to patients conversations. This is to improve overall healthcare for staff and patients.

2.2 What are the benefits of using or sharing the data?

The technology helps healthcare users by streamlining the creation and management of medical records. Automated transcription from patient conversations enables healthcare professionals to focus more on patient care and ensures accurate record-keeping.

3 What data do you want to use or share?

3.1 Can you use anonymous data for your purposes? If not, explain why.

☐	Yes
☒	No
☐	Unsure

The data cannot be anonymous at source as it is collected directly from free form patient conversations. Even if named entities are removed there is no guarantee that multiple pieces of information can be used to identify a person uniquely.

3.2 Which types of personal data do you need to use and why?

Tandem processes personal data to create the basic foundation of medical records.

Sensitive Personal Data: Patient Stories. Information about the patient may include medical history, social situation, symptoms, family situation, sexual habits, alcohol and smoking habits.

Name, address, social security number and other personal data are only processed if they are mentioned during the discussions.

☒	Forename	☒	Physical description, for example height	☐	Photograph / picture of people
☒	Surname	☒	Phone number	☐	Location data e.g. IP address

				□ □	
☒	Address	□	Email address	☒	Audio recordings
☒	Postcode full	☒	GP details	□	Video recordings
☒	Postcode partial	□	Legal representative name (personal representative)	□	Other [see below]
☒	Date of birth	☒	NHS number	□	None
☒	Age	☒	National insurance number		
☒	Gender	☒	Other numerical identifier – Social Security Number		

3.3 Data protection laws mean that some data is considered particularly sensitive. This is called special category data. Data that relates to criminal offences is also considered particularly sensitive. Which types of sensitive data do you need to use or share?

Type of data		**Reason why this is needed (leave blank if not applicable)**
☒	Information relating to an individual’s physical or mental health or condition, for example information from health and care records	As Tandem is used in a medical setting it may process special category data when discussed.
□	Biometric information in order to uniquely identify an individual, for example facial recognition	
□	Genetic data, for example details about a DNA sample taken as part of a genetic clinical service	

☒	Information relating to an individual’s sexual life or sexual orientation	As Tandem is used in a medical setting it may processes special category data when discussed.
☒	Racial or ethnic origin	As Tandem is used in a medical setting it may processes special category data when discussed.
☐	Political opinions	
☐	Religious or philosophical beliefs	
☐	Trade union membership	
☐	Information relating to criminal or suspected criminal offences	
☐	None of the above	

3.4 Who are the individuals that can be identified from the data?

☒	Patients or service users
☒	Carers
☒	Staff
☐	Wider workforce
☐	Visitors
☐	Members of the public
☐	Other

3.5 Where will your data come from?

Directly from the patient's conversation in health care settings.

3.6 Will you be linking any data together?

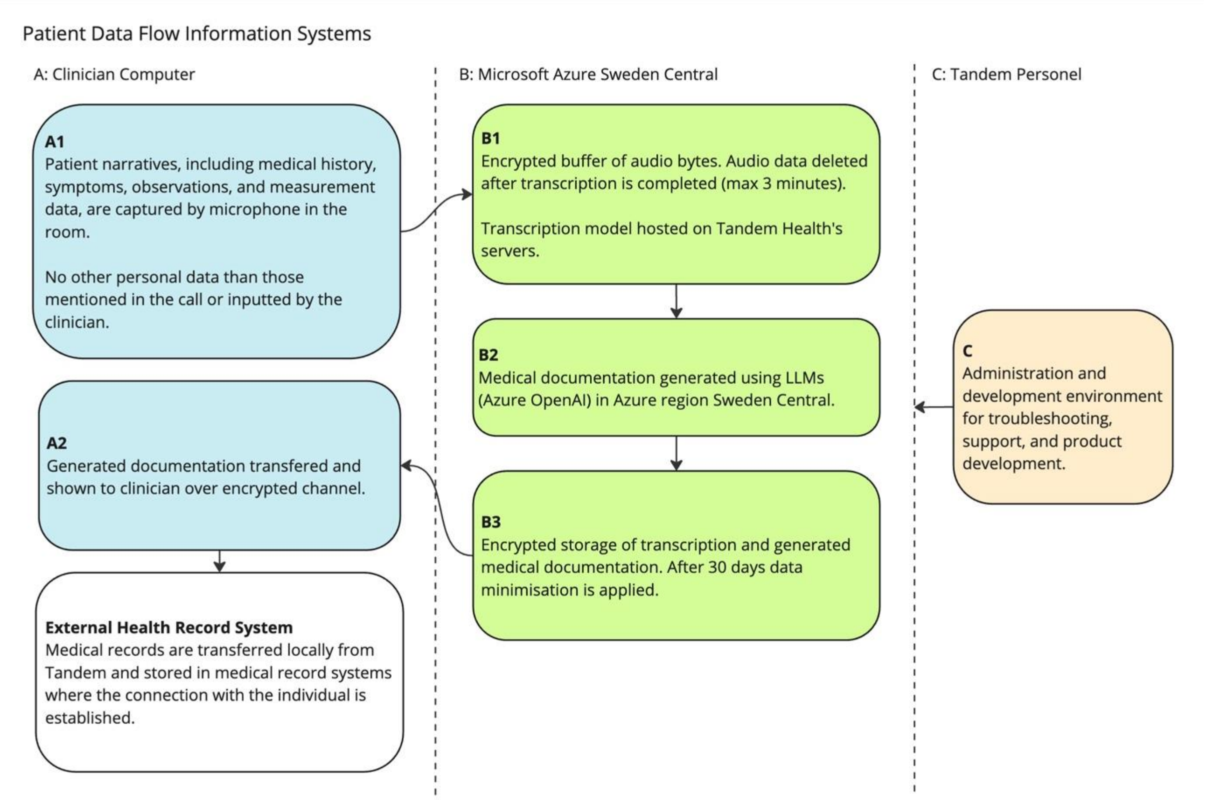
☐	Yes
☒	No
☐	Unsure

3.7 Will it become possible, as a result of linking data, to be able to identify individuals who were not already identifiable from the original dataset?

☐	Yes
☒	No
☐	Unsure

4 Where will data flow?

4.1 Describe the flows of data.



Patient Data Flow Information Systems

A: Clinician Computer

A1
Patient narratives, including medical history, symptoms, observations, and measurement data, are captured by microphone in the room.

No other personal data than those mentioned in the call or inputted by the clinician.

A2
Generated documentation transfered and shown to clinician over encrypted channel.

External Health Record System
Medical records are transferred locally from Tandem and stored in medical record systems where the connection with the individual is established.

B: Microsoft Azure Sweden Central

B1
Encrypted buffer of audio bytes. Audio data deleted after transcription is completed (max 3 minutes).

Transcription model hosted on Tandem Health's servers.

B2
Medical documentation generated using LLMs (Azure OpenAI) in Azure region Sweden Central.

B3
Encrypted storage of transcription and generated medical documentation. After 30 days data minimisation is applied.

C: Tandem Personel

C
Administration and development environment for troubleshooting, support, and product development.

4.2 Confirm that your organisation's information asset register (IAR), record of processing activities (ROPA) or your combined information assets and flows register (IAFR) has been updated with the flows described above.

☒	Yes
☐	No
☐	Unsure

4.3 Will any data be shared outside of the UK?

☒	Yes
☐	No
☐	Unsure

5 Is the intended use of the data lawful?

5.1 Under Article 6 of the UK General Data Protection Regulation (UK GDPR) what is your lawful basis for processing personal data?

☐	(a) We have consent [this must be freely given, specific, informed and unambiguous. It is not appropriate to rely on consent for individual care or research, even if you have obtained consent for other reasons, but is likely to be needed for the use of cookies on a website]
☐	(b) We have a contractual obligation [between a person and a service, such as a service user and privately funded care home]
☒	(c) We have a legal obligation [the law requires us to do this, for example where NHS England or the courts use their powers to require the data. See this list for the most likely laws that apply when using and sharing information in health and care.]
☐	(e) We need it to perform a public task [a public body, such as an NHS organisation or Care Quality Commission (CQC) registered social care organisation, is required to undertake particular activities. See this list for the most likely laws that apply when using and sharing information in health and care. This is mostly likely to be relevant for the provision of NHS and social care services regulated by the CQC. See HRA guidance on legal basis for processing data for research]
☐	(f) We have a legitimate interest [for example, a private care provider making attempts to resolve an outstanding debt for one of its service users. This cannot be relied on by public bodies in the performance of their tasks.]

5.2 If you have indicated in question 3.3 that you are using special category data, what is your lawful basis under Article 9 of the UK GDPR?

☐	(b) **We need it to comply with our legal obligations for employment**
☐	(f) **We need it for legal claims, to seek legal advice or judicial acts**
☐	(g) **We need to comply with our legal obligations to provide information where there is a substantial public interest, as set out in this list**
☒	(h) **We need it to comply with our legal obligations to provide or manage health or social care services**
☐	(i) **We need it to comply with our legal obligations for public health**
☐	(j) **We need it for archiving, research and statistics where this is in the public interest**

5.3 What is your legal basis for using and sharing this health and care data under the common law duty of confidentiality?

☒	**Implied consent** [for individual care or local clinical or care audits. Skip to question 16]
☐	**Explicit consent** [a very clear and specific statement of consent. Go to question 15a]
☐	**Section 251 support** [this means you have support from the Secretary of State for Health and Care or the HRA following an application to the Confidentiality Advisory Group (CAG). CAG must be satisfied that it isn't possible or practical to seek consent. Go to question 15a]
☒	**Legal requirement** [this includes where NHS England has directed an organisation to share the data using its legal powers. State the legal requirement in the further information section. Go to question 15a]
☐	**Overriding public interest** [for example to prevent or detect a serious crime or to prevent serious harm to another person. The justification to disclose must be balanced against the public interest in maintaining public confidence in health and care services. Routine use of this is extremely rare in health and care, as it usually applies to individual cases where decisions are made to share data. Go to question 15a]
☐	**Not applicable** [you are not proposing to use identifiable health and care data. Skip to question 16]

5.3.1 Please provide further information or evidence.

The use of the program is covered by The Common Law Duty of Confidentiality, Health and Social Care Act 2012 based on implied consent, as it is a part of providing direct care to the patient.

Patients have a reasonable expectation that their interactions with healthcare professionals will be documented as part of their medical records to ensure safe, effective, and continuous care. The transcription process and note generation occur within this context and serve solely to assist healthcare providers in maintaining accurate clinical documentation.

6 How are you keeping the data secure?

6.1 Are you collecting information?

☒	Yes
☐	No

6.2 How is the data being collected?

Patient narratives, including medical history, symptoms, observations, and measurement data, are captured by listening in the room.

6.3 Are you storing information?

☒	Yes (however, no audio is stored)
☐	No

6.3.1 How will information be stored?

Storage location	
☐	Physical storage, for example filing cabinets, archive rooms etc
☐	Local organisation servers
☒	Cloud storage

6.4 Are you transferring information?

☒	Yes
☐	No

6.5 How will information be transferred?

Information will be transferred to the electronic medical record system via a "smart copy and paste" feature, ensuring the correct text is entered into the appropriate fields after user approval. Records are transferred locally from the web application to the medical record system.

6.6 How will you ensure that information is safe and secure?

Security measure		**Details (leave blank if not applicable)**
☒	Encryption	The data is encrypted and secure when stored will be taken (following best practice as an ISO27001 certified organisation).
☒	Password protection	
☒	Role based access controls (RBAC)	Access to patient identifiable data will be strictly limited
☒	Restricted physical access	In line with ISO27001

☒	Business continuity plans	In line with ISO27001
☒	Security policies	In line with ISO27001

6.7 How will you ensure the information will not be used for any other purposes beyond those set out in question 2.1?

Specify the measures below which will be used to limit the purposes the data is used for.

Security measure		Details (leave blank if not applicable)
☒	Contract	
☒	Data processing agreement(s)	
☐	Data sharing and processing agreement (DSPA)	
☒	Audit	
☒	Staff training	

7 How long are you keeping the data and what will happen to it after that time?

7.1 How long are you planning to use the data for?

The transcript and related changes are stored at Tandem for as long as the organisation is a customer. Then, all the data, including the logs, will be deleted within 30 days, even from the backup system. Tandem stores data for troubleshooting, support, and product development. This is done as far as possible on data that does not include a name, social security number or address.

7.2 How long do you intend to keep the data?

The transcript and related changes are stored at Tandem for as long as the organisation is a customer. Then, all the data, including the logs, will be deleted within 30 days, even from the backup system. Tandem stores data for troubleshooting, support, and product development. This is done as far as possible on data that does not include a name, social security number or address.

7.3 What will happen to the data at the end of this period?

Action		**Details (leave blank if not applicable)**
☒	Secure destruction (for example by shredding paper records or wiping hard drives with evidence of a certificate of destruction)	
☐	Permanent preservation by transferring the data to a Place of Deposit run by the National Archives	
☐	Transfer to another organisation	
☐	Extension to retention period	
☐	It will be anonymised and kept	
☐	The controller(s) will manage as it is held by them	
☐	Other	

8 How are people's rights and choices being met?

8.1 How will you comply with the following individual rights (where they apply)?

Individual right	**How you will comply (or state *not applicable* if the right does not apply)**	
The right to be informed The right to be informed about the collection and use of personal data.		The data subjects are informed of the processing of their personal data. Tandem informs approved healthcare professionals and healthcare establishments of the treatments it implements and the need for them to: • inform the data subjects (patients) of the processing they implement in their capacity as data controllers using the Tandem solution. • obtain their consent to the processing of their personal data, if applicable.
	☒	Privacy notice(s) for all relevant organisations
	☐	Information leaflets
	☐	Posters
	☐	Letters

	□	Emails
	□	Texts
	□	Social media campaign
	□	DPIA published (best practice rather than requirement)
	□	Other
	□	Not applicable
The right of access The right to access details of data use and receive a copy of their personal information - this is commonly referred to as a subject access request.	Tandem assists by forwarding any requests from data subjects to the Data Controller as soon as possible, in accordance with the instructions set out in the DPAs established between Tandem and all its customers. Tandem, as a processor, may not disclose, suppress or restrict the processing of data to the patient independently without instructions from the controller. At the request of the data controller, healthcare professionals can retrieve support data for up to 30 days after the call, depending on the day and time of the call. It is not possible to search by name or social security number, as this information is not explicitly stored in the medium. Healthcare professionals can copy and share the data in the system with the patient. However, this is mainly done by transferring the data to the medical records system, where patients have access to the notes in the usual order.	
The right to rectification The right to have inaccurate personal data rectified or completed if it is incomplete.	See above regarding Tandem's overall role as a processor and not as a data controller. Specifically, it can be considered that the correction of patient information does not need to be done in the medium, as it is not the primary record used and linked to the patient. Physicians can always choose to adjust/correct the grade before transferring it to the medical record system.	

The right to erasure The right to have personal data erased, if applicable.	A patient can request that their conversations stored at Tandem be deleted or rectified. Such a request is made and executed by the attending physician, in consultation with Tandem if necessary.
The right to restrict processing The right to limit how their data is used, if applicable.	These rights must be guaranteed and implemented by the data controller, i.e. the doctor and/or healthcare establishment that uses the platform.
The right to data portability The right to obtain and re-use their personal data, if applicable.	Patients do not have a right to portability as they have no contact with Tandem. This right can only be exercised by the patient towards his doctor and/or the health care establishment.
The right to object The right to object to the use and sharing of personal data, if applicable.	Again, this is managed by the data controller, i.e. the healthcare organisation using Tandem.

8.2 Will the national data opt-out need to be applied?

☐	Yes
☒	No
☐	Unsure

8.3 Will any decisions be made in a purely automated way without any human involvement (automated decision making)?

☐	Yes
☒	No
☐	Unsure

8.4 Detail any stakeholder consultation that has taken place (if applicable).

Not applicable.

9 Which organisations are involved?

9.1 List the organisation(s) that will decide why and how the data is being used and shared (controllers).

The healthcare organisation using Tandem is the data controller and they decide:

- to collect the data in the first place
- what data is being collected
- what it is being used for
- who it is being collected from

The controller will have a direct relationship with the data subjects from whom the data is being collected, for example patients, service users or employees. For all patient data, Tandem Health remains the processor.

9.2 List the organisation(s) that are being instructed to use or share the data (processors).

Tandem acts under instructions from those listed in question 9.1, for example they are likely to be told:

- what data to collect
- who to collect data from
- how the collection is legal
- the purpose for the collection
- who to share the data with
- how long to keep the data

9.3 List any organisations that have been subcontracted by your processor to handle data

Name of the subcontractor	Purpose	Contract reference	Section 28 Compliance
Microsoft Ireland Operations, Ltd.	Hosting service	September 2023	YES

9.4 Explain the relationship between the organisations set out in questions 28, 29 and 30 and what activities they do

Microsoft Ireland Operations, Ltd host the servers and databases used for the processing and storing.

9.5 What due diligence measures and checks have been carried out on any processors used?

Due diligence measures	
☒	Data Security and Protection Toolkit (DSPT) compliance
☒	Registered with the Information Commissioner's Office (ICO)
☒	Digital Technology Assessment Criteria (DTAC) assessment
☒	Stated accreditations
☒	Cyber Essentials or any other cyber security certification
☒	Data Security and Protection Toolkit (DSPT) compliance

10 What data protections are there and what mitigations will you put in place?

10.1 Complete the risk assessment table. Use the risk scoring table to decide on the risk score.

Risk assessment table

Risk ref no.	Description	Risk score* (L x I)	Mitigations	Risk score* with mitigations applied
R-2	***The patient is not informed of the use of Tandem.***	Likelihood: 4 Impact: 1 **Risk Score: 4**	All users are given a clear introduction on informing the patients before using Tandem.	Likelihood: 3 Impact: 1 **Risk Score: 3**
R-5	***Sensitive patient information is leaked from language model***	Likelihood: 1 Impact: 4 **Risk Score: 4**	A specific agreement with the language model provider not to store information sent to them. Only servers based in EU, are used.	Likelihood: 1 Impact: 4 **Risk Score: 4**
R-9	***The data minimization algorithm fails and data containing name, social security number, or address is stored over time at Tandem Health.***	Likelihood: 2 Impact: 2 **Risk Score: 4**	The data minimization algorithm is continuously developed and its performance is monitored in line with ISO27001. Continued use of Microsoft's most powerful solution for this type of task.	Likelihood: 1 Impact: 2 **Risk Score: 2**
R-8	***Medical information accessed by unauthorised persons on user's device***	Likelihood: 3 Impact: 2 **Risk Score: 6**	Automatic logout after inactivity, routines to always log out if you leave the room for a long time. The login systems comply with the highest standards with strong authentication with multi factor authentication (MFA).	Likelihood: 2 Impact: 2 **Risk Score: 4**
R-14	***Systems are unavailable due to unforeseen capacity constraints.***	Likelihood: 2 Impact: 3 **Risk Score: 6**	Systems are maintained, deployed and monitored in line with ISO27001.	Likelihood: 1 Impact: 3 **Risk Score: 3**
R-16	***Incident response is slow and ineffective.***	Likelihood: 2 Impact: 3 **Risk Score: 6**	Systems are maintained, deployed and monitored in line with ISO27001.	Likelihood: 1 Impact: 2 **Risk Score: 2**

R-17	***Company systems and/or data are breached via a compromised development environment or system.***	Likelihood: 2 Impact: 3 **Risk Score: 6**	Development, test and production environments are strictly separated. As such a breach of test or development won't result in a breach of sensitive production data.	Likelihood: 2 Impact: 2 **Risk Score: 4**
R-10	***Attackers gain unauthorized access to employee/admin/developer accounts.***	Likelihood: 2 Impact: 4 **Risk Score: 8**	Internal training on security procedures. MFA for all staff. Recording access to sensitive data. Monitoring systems on Azure, EDR for developers. RBAC based on principle of least privilege. Regular access reviews.	Likelihood: 1 Impact: 3 **Risk Score: 3**
R-13	***Vulnerabilities are introduced by developers resulting in a compromise of company systems and/or data.***	Likelihood: 2 Impact: 4 **Risk Score: 8**	Internal training on security procedures. Secure coding practices and quality control in line with ISO27001. Github Dependabot for detecting and mediating external vulnerabilities.	Likelihood: 1 Impact: 4 **Risk Score: 4**
R-18	***Sensitive data is breached in transit due to improper encryption.***	Likelihood: 2 Impact: 4 **Risk Score: 8**	Data is encrypted during transit using SSL. Software is developed according to quality control best practices in line with ISO27001.	Likelihood: 1 Impact: 4 **Risk Score: 4**
R-19	***Company systems and data are breached by a company vendor.***	Likelihood: 2 Impact: 4 **Risk Score: 8**	Any new vendors are reviewed when collaboration is initiated and yearly following that. Critical & high-risk vendors need to follow the highest security standards.	Likelihood: 1 Impact: 4 **Risk Score: 4**
R-20	***Company data is breached, corrupted or made unavailable due to a malware attack.***	Likelihood: 2 Impact: 4 **Risk Score: 8**	Code dependencies analysed by Github Dependabot. Container dependencies analysed using Microsoft Defender for Cloud. VPN requirements for accessing production systems. EDR installed on critical administrator devices.	Likelihood: 1 Impact: 4 **Risk Score: 4**

***Risk scoring table**

		Impact (I)				
		Negligible (1)	**Low (2)**	**Moderate (3)**	**Significant (4)**	**Catastrophic (5)**
Likelihood (L)	**Rare (1)**	**1**	**2**	**3**	**4**	**5**
	Unlikely (2)	**2**	**4**	**6**	**8**	**10**
	Possible (3)	**3**	**6**	**9**	**12**	**15**
	Likely (4)	**4**	**8**	**12**	**16**	**20**
	Almost certain (5)	**5**	**10**	**15**	**20**	**25**

1. Additional Measures

In view of the probability and severity of the risks mapped as well as the technical and organisational measures put in place, it is considered that the risks are limited.

In order to ensure that the persons concerned by the processing carried out by health professionals and health establishments are properly informed, Tandem provides a standard document for the latter in order to facilitate the fulfilment of their obligation to inform.

In addition, as the end users of the Support (healthcare professionals and healthcare establishments) have a key role in the protection of the personal data of the persons concerned, Tandem provides a list of good practices in order to make them aware of the issues in this area.

11 Review and sign-off

Reviewer sign-off	
Reviewer name:	Gustaf Johnssén
Reviewer job title:	DPO
Reviewer contact details:	gustaf.johnssen@tandemhealth.ai
Date of review:	2025/01/24
Comments:	
Date for next review:	

Approver sign-off	
Approver name:	Oliver Åstrand
Approver job title:	CTO
Approver contact details:	oliver.astrand@tandemhealth.ai
Date of approval:	2025/01/27
Comments:	