



Framework for Data Ethics Tenets in Healthcare with Use Case

Kavita Khobragade¹(✉), Poonam Ponde², and Manisha Bharambe³

¹ Fergusson College (Autonomous), Pune, India

kavita.khobragade@fergusson.edu

² Nowrosjee Wadia College (Autonomous), Pune, India

poonamponde@nowrosjeewadacollege.edu.in

³ MES Abasaheb Garware College (Autonomous), Pune, India

mgb.agc@mespune.in

Abstract. Innovation in latest technologies have provided means to gather data using various ways. Almost all the domains generate, store, and analyze the data for the improvement of the services they provide. The healthcare industry too generates a significant data which is used for improving public healthcare. While dealing with health data, it is necessary to follow appropriate data ethics as health data is considered the most sensitive and it needs to be properly collected, stored, processed and shared with different domains. This research paper discusses about the various data ethics to be followed to handle individual's health data, suggests a framework to deal with this data and a use case is suggested to understand the data ethics tenets.

Keywords: Healthcare Data · Data Source · Data Ethics · Tenets · Accountability · Privacy · Fairness · Transparency

1 Introduction

Today, we live in a world which is data driven world. Huge volumes of data is generated by applications in all domains including healthcare. Organizations such as hospitals, health centers, municipal corporations, insurance companies and others collect health related data for various purposes. Wearable fitness devices and apps also collect personal health data on a daily basis. The government health agencies and related departments track and collect health information of the general public using various methods such as Aarogya Setu app. The World Health Organization provides tools which helps countries to strengthen the capacity of collecting, compiling, managing, analyzing and using health data.

With the rise in the use of tools and technologies, there has been immense digital transformation in the healthcare systems all around the world. Introduction of the Web, Clouds, Big Data, Internet of Things and various other Smart applications in almost every aspect of our lives, have contributed towards improving our lives enabling us to do otherwise impossible tasks [1]. This Digital transformation has led to collection of

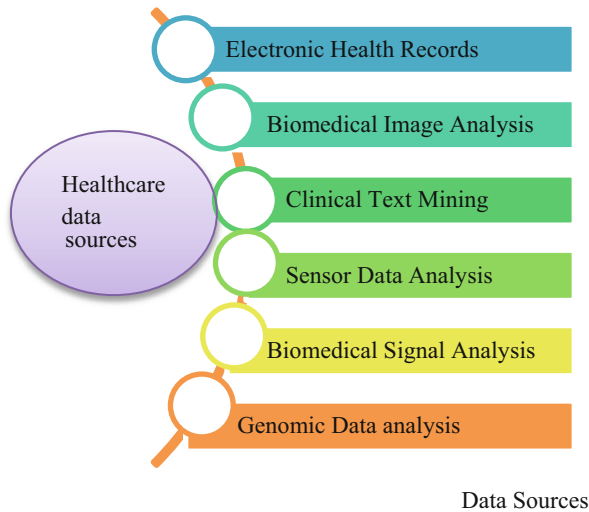


Fig. 1. Healthcare Data Sources

plethora of data that can be useful for the betterment of healthcare all over the globe. This data can be efficiently utilized for the benefit of the people and can also be misused. That is why data ethics becomes one of the important factors to protect the sensitive healthcare data.

1.1 What is Healthcare Data?

Healthcare system aims at providing facilities to diagnose, treat and prevent various health-related issues. This system generates huge amount of data every day that is useful as well as challenging for development of better health solutions. This data is of different types like patient's information, medical history, medical and clinical data and other private or personal medical data [7].

1.2 Healthcare Data Sources

The healthcare data can be collected from various sources as shown in Fig. 1. This healthcare data is rich and it is derived from a wide variety of sources such as images, sensors, text, electronic records etc. [8].

1.2.1 Electronics Health Records

The EHR i.e. Electronics Health Record contains a digitized version of patient's medical history such as patient's demographic data, problems, medications, physician's observations, medical history, laboratory data, radiology reports, progress note and billing data, etc. [8]. EMR [Electronic Medical Record] or CPR [Computer-based Patient Record] are the synonyms used for representing EHR. The first known medical record was suggested in fifth century by Hippocrates, and they prescribed two goals of medical records: i) The

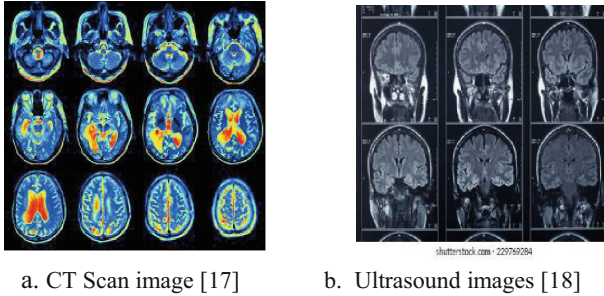


Fig. 2. Biomedical Images of Human Brain [CT Scan & Ultrasound]

medical record should accurately reflect the course of disease and ii) It should indicate the probable cause of disease [11].

1.2.2 Biomedical Image Analysis

Biomedical imaging plays a very important role to provide high-quality images of anatomical structures of human being. There are many modalities by using which medical images can be collected. Some of the modalities are MRI [Magnetic Resonance Imaging] which examines organs, tissues and skeletal system, CT-Scan [Computer Tomography] which is used to diagnose bone tumors and fractures, PET [Positron Emission Tomography] used to identify cancer, heart diseases and brain disorders and Ultrasound imaging which uses the sound wave to produce images of body structures. These technologies help physicians to look inside the body of a patient without hurting the patient and allow to better understand the cause of an illness [11].

1.2.3 Clinical Text Mining

Most of the health records of a patient is maintained as clinical notes in the hospital/clinic which are generally in unstructured format. Clinical notes contain the clinical information such as medication prescribed, percentage of doses in injection, saline doses etc. The clinical notes vary across different patients and practitioners [8]. Storing, analyzing and handling this unstructured data becomes very difficult. As it is textual data, handwriting matters a lot when it is supposed to be written and stored on digital media.

1.2.4 Sensor Data Analysis

The real time health data can be collected by using sensors. Electrocardiogram [ECG] and Electroencephalogram [EEG] are the examples of sensors which collects signals from different parts of the human body. This data is useful in ICU's [Intensive Care Units] to analyses the condition of a patient and his treatment [10]. This is shown in Fig. 3.

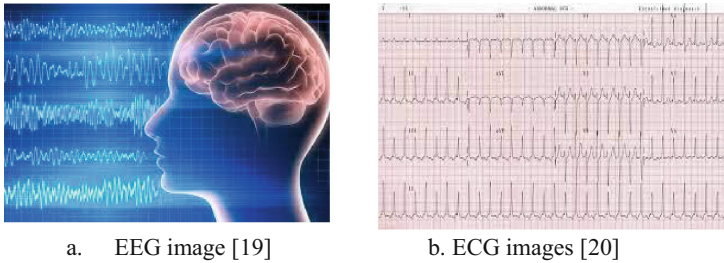


Fig. 3. Sensor Data Analysis of Human Brain [EEG & ECG]



Fig. 4. Biomedical Signal Analysis [PEG image] of Human Brain [21]

1.2.5 Biomedical Signal Analysis

The signals from human body are collected by using various sensors listed here. Biomedical signal analysis basically processes these signals. It includes Electrocardiogram [ECG], Electroencephalogram [EEG], Electroneurogram [ENG], Electromyogram [EMG], Electrogastrogram [EGG] and Phonocardiogram [PCG]. The analysis of these signals helps physician to decide the correct pathway for treatment [8] (Fig. 4).

1.2.6 Genomic Data Analysis

Genomic data analysis deals with diseases that are caused due to genetic factors [8]. While acquiring the data from various sources above, the data ethics should be followed which is helpful for storage, processing and sharing this data.

2 What is Data Ethics?

Data ethics can be defined as the branch of ethics which studies and evaluates moral problems related to data such as collection, processing, dissemination, sharing, using, etc. [12]. Data ethics is implicitly a set of commandments about the usage of data i.e. what you may or may not do with the data. In other words, data ethics is committing to consider the ethical consequences of your behavior [9].

2.1 Ethical Issues in Healthcare Data

While working in the Healthcare domain, the data can be collected in two categories for research activities such as data of unaffected People (i.e. Common People) and data of affected People (i.e. Patients).

The flow of the complete process would include collection the data, storage of the data, analyze or process the data and share the data. While going through these steps there are certain aspects which are very important for maintaining the patient's autonomy, data protection is must. The patient's data must not be used without their consent as there is a potential risk of disclosing patient's information. In order to avoid this loss of autonomy, quality assurance activities must be ethically strong. While going through the process, few measures must be considered about the data such as accountability, security, privacy, fairness, respect for a person and transparency [13]. These attributes are discussed in detail below:

Accountability

Mechanisms of storing the data must ensure the accountability to maintain the fair and transparent processing of personal data. This will be helpful in order to protect the human biological information. This will lead to quality assurance of the resource tools [14]. Organization's reasonable and systematic accountability is to use and protect data in a systematic way spontaneously and reasonably. The risk of individual data usage is reduced by accountability. The social and ethical implications are also reduced.

Data Security

With the amount of healthcare data getting stored on cloud as well as being shared between individuals and organizations online, the vulnerability to data has increased manifold. Data security implies controlling the access to information and also allowing easy access to those who need that information for a good purpose. Data security mainly means protecting the data against unauthorized access. It also concentrates on protecting data from destructive attacks and stealing data for profit and gains.

Recently there have been numerous security breaches into healthcare data leading to misuse of the individual data like the Blackbaud Ransomware Attack [15] or the recent cyber-attack on a Finnish psychotherapy provider [2] where the sensitive data of the patients was stolen and they were threatened for ransom.

Due to the sensitive nature of healthcare data, it is the priority to make the data security services very reliable and robust. A reliable way for authentication of users of the data when gaining access to the patients' data and records must be adopted. Proper data protection strategies should be implemented and continuous monitoring and reporting of regulations and measures should be followed to keep the data secure.

Privacy of Data

In recent times, there has been a tremendous increase in the use of digital technologies to gather and store data about health globally. This data may be available on various online platforms like servers, mobile phones etc. Therefore, maintaining the privacy of the users becomes the most vital activity [3]. It becomes very necessary to safeguard the privacy of the people participating in sharing their health information. The participant should be made fully aware of whether their data will be shared with third parties. The organizations or researchers gathering the data should gain the trust of the participants that their sensitive health data will not be distributed or used without their authorization and consent. Because loss of privacy may result in unexpected side effects like emotional distress or embarrassment and discomfort. For example, for studies related to diseases

like AIDS or even the current pandemic of COVID-19, leak in the privacy of data may lead to social ostracizing, discrimination and denial of services. To avoid this, procedures to maintain the anonymity and confidentiality of the participants should be followed [4].

Sometimes it may be essential to share the data with third-party organizations for the benefit of the society and developing the overall health services provided to the public. In this case, it should be important to build a good data sharing environment where the health data will be shared for the betterment of the society along with maintaining the individual's right to privacy and consent [16].

Fairness

Fairness of data means that the data that has been gathered should be accessible and reusable for further research purposes for getting more advanced knowledge. In the field of healthcare, historic data and the previous results of that data plays a vital role in developing more advanced cures and medicines for certain health problems. That is why this data should be easily accessible to the authorized and authenticate individuals or organizations [14].

While maintaining this fairness, it is also important to safeguard the individual's privacy and confidentiality. If vital information is to be shared for better outcome, then the right to consent of the person should also be considered.

Respect for a Person

It is important to maintain the rights and dignity any of the individual sharing their personal data for the health care and research activities [14]. In this process, respect of a person and caring about the data matters a lot. It should be maintained with the highest priority. While storing the data all 6 above mentioned measures should be considered equally and all of them are related with each other.

Transparency

In this process, patient or Common Person both has full right for accessing the collected and stored data. Both the individuals should have the right to know about any kind of processing activities where their personal data is being used. Activities may include the National level research programs. This refers to the transparency of the data. It should be maintained throughout the processes. Transparency about the storing and processing of data will help to increase the individual's trust in these processing activities. It will ensure more safe resources as transparency will allow individual to verify the conclusions drawn and correct the mistakes too [13]. The efforts should be made in a way such that it will lead towards minimizing the risk which is caused due to travelling of data in the data supply chain which can be archived by increasing the transparency while collecting and storing the data [5].

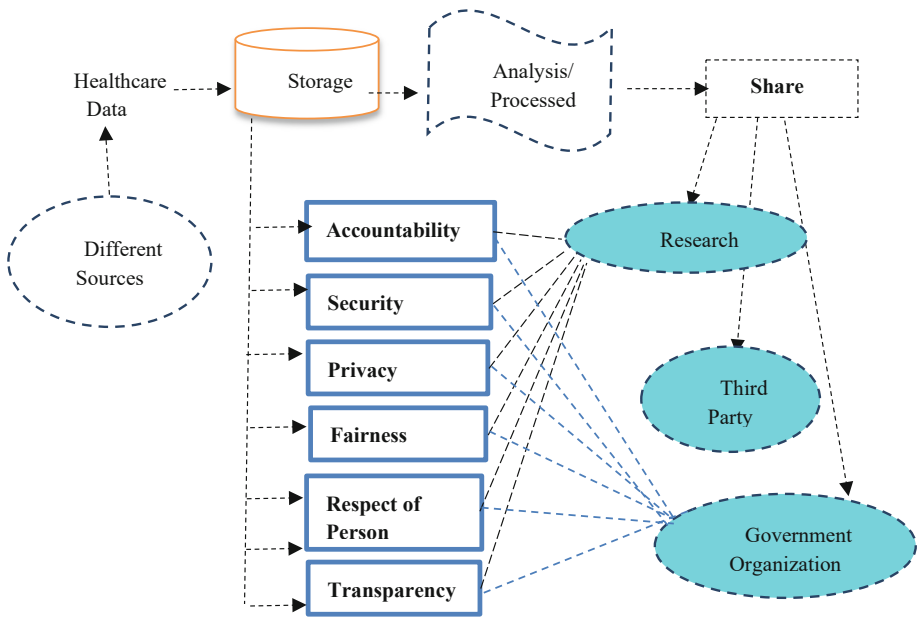


Fig. 5. Framework of Ethics in Healthcare [10]

3 Proposed Framework for Ethics in Healthcare

After considering all the above aspects required for the any Data Ethics Frameworks, following framework is proposed for the Healthcare Research area. The proposed framework for ethics is shown in Fig. 5 [10].

In this framework, the healthcare data of a common person or a patient can be collected from various sources. This collected data is treated as healthcare data and it is stored into some storage location or device. This data can be used and shared by different groups like researchers, third parties or any organization for advanced research activities or for improving various healthcare programs. While storing, processing and sharing this data, the above enlisted aspects, which in turn are called as ethics, should be followed by each of these groups.

When the data is shared with researchers or third party or with any organization, all the ethics should be followed while preserving the sensitivity of the data.

3.1 Hardware and Software Testbed for Framework

This paper is based on the approach of frontend, middleware and backend. As it is a framework, there are various issues need to be considered while generating testbed. These issues are volume of data, variety of data from different sources, data security

and availability of data. The medical data is collected from various sources. The data collected and stored is of huge volume and continuously increasing. Rather than investing in physical servers for data storage, the cloud storage (for example- AWS Cloud or Google Cloud) can be used which will take care of security and scalability of medical data. The users who wish to access data, use some user interface (for example- mobile user interface) and access the data. This data access is provided through firewall and VPN for secure access across public networks like internet. A middleware is used to provide access rights to different users and authorities with different access rights. Due to this, the data is shared with the correct person.

A typical testbed for the above framework is as follows:

Web Server – Apache

Database – MS SQL

Operating System - Windows/Android

Browser – Firefox

Frontend – Javascript/HTML/CSS

3.2 Use Case for ensuring Data Ethics Tenets

The Use Case provides the example of ethical consideration with data encountered while handling healthcare framework. A use case is a method to show the open-ended scenarios which helps readers to observe the issues and ethical considerations in each phase of data lifecycle. The questions raised in a use case are useful for healthcare agency for collecting, analyzing and posting the data. The ethical data usage is best to show the ethical behavior of the healthcare agency or organization. The Framework's Use Case aims to complement existing cultural norms and training initiatives that reinforce ethical behavior across the healthcare agency [5, 6].

3.3 Use Case: Dissemination and Impacts

Dissemination is the action or fact of distributing, spreading, broadcasting some information widely. Due to the dissemination what impact occurs on healthcare data and its use case is shown in Table 1.

Table 1. Use Case: Dissemination and Impacts

Dissemination and Impacts	
Organization Type:	Government Health Organization Agency
Primary Objective:	Enforce the Law and minimum standards of data ethics in healthcare
Secondary Objective:	Publish report to increase the public understanding on data ethics in healthcare
Scenario:	<p>The data ethics in healthcare regulates and maintains the security, privacy and fairness of the patient data, which can be used in research, shared with third party for commercial sale or exhibited to public via government organizations. The government health organization agency enforces the law with minimum standards using which the healthcare data can be assessed. Each year, the agency collects, manages, and analyzes healthcare records of common persons and patients. The reports can be generated and the results can be posted on website. Consider a scenario of pandemic. The pandemic related data was posted on website. Few people were abused and they raised complaint for the same. The public and activists were not happy with this scenario. They demanded for public transparency and respect to a person. They wanted to know who should be held accountable for the situation. Due to the privacy concerns with public demand, the healthcare agency removed certain reports from the website. For few cases, in place of removing the reports from website, they have taken some measures to redact the healthcare information previously posted. This leads the healthcare agency and its team into conflicting views about how to proceed.</p>
General Use Case Questions:	<p>The use case questions can be raised for checking the ethics in healthcare. These are:</p> <ol style="list-style-type: none">1. Should healthcare agency remove some report or records from website rather than revising the personal and private data? (For promoting transparency)2. Is there any method available to prevent such situations? (For promoting Accountability)3. If vital information is to be shared for better outcome, then does the healthcare agency takes consent of the person? (For promoting Right to Person)4. Is there any specific policy or procedure available which will provide guidance to people for such situations? (For promoting Privacy and Security)

(continued)

Table 1. (continued)

Data Life Cycle Questions:	<p>Healthcare Data Collection (Acquisition)</p> <ol style="list-style-type: none"> 1. Does the healthcare agency have any right for collecting and releasing healthcare data? (Who has collected the data matters a lot in healthcare) 2. Is there any such document already exists? If exists, does it need to be updated? (For privacy and security reasons) 3. Is there any data quality check performed on the collected data? 4. Who held accountable for collecting data? 5. After data collection, does any preprocessing done on the data before it gets released? 6. While analyzing data, what privacy rules are applied? 7. What healthcare data need to be shared for public? 8. Who will held accountable if data consumers use the data in a harmful way? 9. Is there any specific right given to consumers for using the data? 10. Could the healthcare data be combined with other data to identify or analyze personal information? 11. How long this data should be maintained? 12. Is there any mechanism available for data storage? Could we make changes in it? Does it provide any impact on long view of data?
-----------------------------------	--

4 Conclusion

In the real world, transparency should be maintained with healthcare data as it is very sensitive data. This data should not be shared amongst different stakeholders. The healthcare data should be privacy preserving data. The different stakeholders such as patients, doctors, third party people, etc., should be adequately informed about the purpose of data captured, how it is collected and stored, transferred or processed. This data should not be shared or sold to any third-party companies or institutions and there should be a guarantee given to the patient. Everyone should understand the legal, ethical and cultural challenges related to healthcare data. These new policies should be defined for the maintenance of healthcare data. A use case is a method to show the open-ended scenarios which helps readers to observe the issues and ethical considerations in each phase of data lifecycle. The Use Case was provided as the example of ethical consideration while handling healthcare framework. The general use case questions and data life cycle questions are the best examples for healthcare officials. The suggested Framework's Use Case is a best method to follow and understand various data ethics and helps to handle individual's health data.

References

1. Rue Montoyer, Healthcare, artificial intelligence, data and ethics - A 2030 vision 51, Brussels, 1000, Belgium. <https://www.digitaleurope.org/wp/wp-content/uploads/2019/02/Healthcare-AI-Data-Ethics-2030-vision.pdf>, <https://www.hipaajournal.com/september-2020-healthcare-data-breach-report-9-7-million-records-compromised/> referred on 7th November 2020 at 3.45pm, (December 2018).
2. MacPherson Y., Pham K., Ethics in Health Data Science. In: Celi L., Majumder M., Ordóñez P., Osorio J., Paik K., Somai M. (eds) Leveraging Data Science for Global Health. Springer, Cham. https://doi.org/10.1007/978-3-030-47994-7_22, (2020).
3. Barrow JM, Brannan GD, Khandhar PB, Research Ethics. [Updated 2020 Sep 25]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; Available from: <https://www.ncbi.nlm.nih.gov/books/NBK459281/>, (Jan 2020).
4. Riso, B., Tupasela, A., Vears, D. F., Felzmann, H., Cockbain, J., Loi, M., Kongsholm, N., Zullo, S., & Rakic, V., Ethical sharing of health data in online platforms - which values should be considered. Life sciences, society and policy, 13(1), 12. <https://doi.org/10.1186/s40504-017-0060-z>, (2017).
5. Victor Chang, Rahman Olamide Eniola, Ben Shaw-Ching Liu, Mitra Arami, An Ethical Framework for Big Data and Smart Healthcare, DOI: <https://doi.org/10.5220/001103090000320>.
6. David J., An Article on Hand on the ethical, social, and policy challenges associated with the rise of “big data,” available at <https://imstat.org/2018/12/14/hand-writing-right-legitimate-and-proper-the-new-world-of-dataethics/>, (2018)
7. Chandan K. Reddy, Charu C. Aggarwal, Healthcare Data Analytics, CRC Press, (2015).
8. Mark A. Musen, Jan Hendrik Bennel, Handbook of Medical Informatics, MIEUR, (1999).
9. Joel Grus., Data Science from Scratch First Principle with Python, O'Reilly Publications, (May 2019).
10. Prof.Dr.K.Muthuchelian, Dr.P.Sundara Pandian, Dr.N.Jeyakumaran, Dr.J.Pandiarajan, Dr.M.Monisha, Glimpses of Engineering and Technology in the Modern World, Virudhunagar Hindu Nadars' Senthikumara Nadar College (Autonomous), ISBN: 978-81-942052-7-2, December, 2020.
11. <https://royalsocietypublishing.org/doi/https://doi.org/10.1098/rsta.2016.0360> referred at 8.00 pm on 05-11-2020, (2020).
12. <https://www.twobirds.com/en/news/articles/2019/global/big-data-and-issues-and-opportunities-transparency-consent-control-and-personal-data-ownership> referred at 2:54 pm on 7 November 2020, (2020).
13. <https://ethicalsystmsmap.hesethz.ch/map> referred at 3:47 pm on 7 November 2020, (2020).
14. <https://www.hipaajournal.com/september-2020-healthcare-data-breach-report-9-7-million-records-compromised/> referred at 3.45pm on 7th November 2020, (2020).
15. <https://www.hipaajournal.com/hackers-blackmail-finnish-psychotherapy-provider-and-patients-and-leak-psychotherapy-notes/> referred at 3.50 pm on 7th November 2020, (2020).
16. https://www.accenture.com/_acnmedia/pdf-22/accenture-data-ethics-pov-web.pdf referred at 1:40 pm on 7th November 2020, (2020).
17. <https://americanhealthimaging.com/blog/brain-scan-help-diagnose-adhd/>
18. <https://www.google.com/search?q=ultrasound%20images%20of%20human%20brain&tbm=isch&hl=en&tbs=ring:CVyOSJBKSxZCYfuRqGsbVAdsgIMCgI-IABAAOgQIARAaWAIA&sa=X&ved=0CB0QuIIBahcKEwiQibDLsa7-AhUAAAAAHQAAAAQFA&biw=1519&bih=688#imgrc=xtOLtPBzFcCp1M>
19. <https://www.news-medical.net/news/20220225/Researchers-study-EEG-patterns-in-dying-man-to-understand-brain-activity-at-the-end-of-life.aspx>

20. https://www.google.com/search?q=ECG+images&tbm=isch&hl=en&chips=q:ecg,g_1:tachycardia:EhwgWG9DDIQ%3D,g_1:supraventricular:EhwgWG9DDIQ%3D&sa=X&ved=2ahUKEwjP6LKltK7-AhXX13MBHTMqAd0Q4lYoBX-oECAEQMg&biw=1519&bih=688#imgrc=SIM5-1BAaSIJEM
21. <https://www.spring.org.uk/2014/10/the-personality-trait-that-doubles-alzheimers-risk.php>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

