Ethics of healthcare robotics: Towards responsible research and innovation

CrossMark

Bernd Carsten Stahl [a,*], Mark Coeckelbergh [a,b]

[a] De Montfort University, United Kingdom
[b] University of Vienna, Austria

HIGHLIGHTS

- Traditional approaches to the ethics of robotics are often distant from innovation practices and contexts of use.
- We list key concerns of ethics of healthcare robots.
- Collaborative and embedded ethics can help address ethics of healthcare robotics.
- Responsible research and innovation (RRI) offers a broad array of tools to ensure acceptability of technology.
- RRI in ICT can point out how social concerns can be incorporated.

ARTICLE INFO

ABSTRACT

How can we best identify, understand, and deal with ethical and societal issues raised by healthcare robotics? This paper argues that next to ethical analysis, classic technology assessment, and philosophical speculation we need forms of reflection, dialogue, and experiment that come, quite literally, much closer to innovation practices and contexts of use. The authors discuss a number of ways how to achieve that. Informed by their experience with "embedded" ethics in technical projects and with various tools and methods of responsible research and innovation, the paper identifies "internal" and "external" forms of dialogical research and innovation, reflections on the possibilities and limitations of these forms of ethical–technological innovation, and explores a number of ways how they can be supported by policy at national and supranational level.

1. Introduction

The past decade has seen a rapid growth of research in the area of ethics of robotics, also and particularly as applied to healthcare. This is unsurprising, since research and innovation in the area of healthcare robotics has seen a significant growth in recent years. Consider for instance research presented in this journal: in response to challenges related to ageing, care robots have been developed to support elderly people living at home (e.g. [1]), robotic nurses have been created to assist with care tasks (e.g. [2]), surgical robots have been designed and used in hospitals (e.g. [3]), and robots have been made more socially interactive (e.g. [4]), which also supports the development and use of robots in health care contexts.

Responding to what is taken to be the near future of health care, ethicists have been especially concerned with what is supposed to be the prospect of intelligent, autonomous, and often also humanoid robots that take care of the elderly. Questions addressed include: Will robots replace the nurses and other care givers, leaving the ill and elderly in the hands of machines? Could robots deliver the same quality of care? Can machines give the "warm", "human" care we seem to expect from human care givers? Do robots used in care deceive vulnerable persons when they (the robots) "pretend" to be something else than they are, for example when they appear as pets (see Section 2)?

These reflections on the future of "machine" healthcare are helpful ways of exploring ethical sensitivities about healthcare, thinking through some of our ethical concerns, developing more refined arguments about what exactly we think might be problematic, and better understanding the current developments in the context of modern healthcare and its politics and the wider developments in robotics and our technological culture. However, they are somewhat limited when it comes to *changing* how things

---

are done in healthcare research, innovation and practice. This is partly so since the context in which academic reflection and research in ethics takes place is largely divorced from the context of innovation and practice. How can this gap be bridged?

In this paper we argue that traditional ways of bridging this gap such as case studies and, more recently, desk-based "value sensitive design"-oriented work, are insufficient to transform ethics of healthcare robotics in a way that really engages with problems as they emerge in innovation, user, and stakeholder contexts. First we review the methods of philosophical reflection on ethical issues, using case studies, and thinking about values in design. Then we discuss what we take to be more dialogical, more democratic, and more effective ways of *doing* ethics: (1) collaborative, "embedded" ethics *in* healthcare robotics which directly and substantially involves ethicists in innovation and user processes and (2) various ways of *really* involving stakeholders in innovation and practice, thus rendering healthcare robotics more ethically and socially responsible. For this purpose we introduce the concept of responsible research and innovation (RRI) and show how a generic framework for RRI in ICT can be applied to healthcare robotics. We frame these options as representing forms of "internal" and "external" dialogue. We also reflect on what the current societal and organisational barriers are that prevent these methods from being widely adopted, and we critically discuss the problems and limitations of these methods. Finally, we reflect on what kind of policies may support these "closer" forms of ethical–technological innovation.

2. Traditional approaches to ethics of healthcare robotics: Philosophical reflection on ethics of healthcare robotics, case studies, and desk-based value sensitive design

There is a growing body of literature on the ethics of healthcare robotics and ICTs [5–12], sometimes also called machine (medical) ethics [13,14,8,15–17]. The literature gives a good overview of potential ethical issues in healthcare robotics and shows that philosophical reflection delivers valuable insights into what exactly might be problematic in this area and why.

2.1. Ethical concerns

Here are some ethical and social issues and philosophical discussions we identify as central. This subsection is not meant to be comprehensive; it is meant as a pragmatic and heuristic tool to gain an overview, before we begin the development of the main arguments of this article.

First, there are critical evaluations of healthcare technology visions in terms of their implications for society and on healthcare, for example:

- **Replacement and its implications for labour**: Are robots introduced to solve problems in healthcare and elderly care, or are they introduced to save money by replacing human care givers by robots, and to help robotics research and industry? For instance, in research concerning the development of robots for the elderly, robots are often presented as a response to demographic challenges (see again [1]). But are such technological solutions the main or only way we should tackle these challenges? And if there is truth in the suspicion that robots will replace humans, which problems exactly would they solve, and is robotics really a threat to employment? More generally, what are the consequences for healthcare work? For example, do robots and ICTs threaten "care craftsmanship" [7]?

- **Replacement and its implications for the quality of care: de-humanisation and "cold" care**. An important fear in discussions about robots in healthcare is that robots may replace human care givers, and that this may not only put these people out of job, but also remove the capacity for "warm", "human" care from the care process. It is highly doubtful, for instance, if robots could ever be empathic [39] or have emotions [18]. Robots, it seems, are not capable of a "human" kind of attention and care, whereas healthcare seems to involve more than some "behaviours"; humans have various social and emotional needs, which are not necessarily met by giving them a robot. "Machine care" sounds cold and mechanical. There is the concern that elderly people are abandoned, handed over to robots [11] devoid of human contact [10]. More generally, do machines in care "objectify" care receivers? Do they objectify *care givers* (see also the previous point)? What do we mean by good healthcare? Do we have good healthcare today, without even considering robots? Is good care possible in the context of modernity [14]?

Second, there are issues that have less to do with the idea of replacement as such but are raised by human–robot interaction in healthcare and especially by the robot taking over tasks from humans, for instance:

- **Autonomy**. Not all health care robots are autonomous robots. For instance, surgical robots are remote controlled by the surgeon. Yet health care research often aims to give more autonomy to the robot. An important term in the field, for instance, is *autonomous systems* (see also the title of this journal). Autonomy means here that the robot is designed to carry out tasks without continuous human guidance and assistance, preferably in an unstructured environment. This development could lead to a future scenario in which robots would replace human care workers, for instance if care robots take over the work of the human nurse. As indicated before, this is ethically problematic. But even if robots in healthcare did not entirely replace human care workers, there is still the question how autonomous (in the sense of doing tasks on its own, unassisted by humans) the robot would be and should be in the context of the interaction and the care, and how autonomous it should be in the sense of operating without human supervision. For example, if robots are used in therapy for children, should the robot be supervised (and if so in what way) and what exactly and how much should it do without direct human intervention? (See for instance [19].)

- **Role and tasks**. Related to the previous point is the question regarding the role of the robot in the particular care process. Even if humans are still part of the care process, what exactly should the role of the robot be (and the role of the human)? What tasks can and should be delegated to robots? And in general: should they assist or take over human tasks? When and where should they do what?

- **Moral agency**. Robots do not seem to have the capacity of moral reasoning or, more generally, of dealing with ethically problematic situations. Hence when a moral problem arises within the human–robot interaction and within the healthcare situation, there seems to be a problem: the robot is given (more) autonomy, in the sense of doing tasks by itself without human intervention, but does not seem to have the capacity of moral agency: it can do all kinds of things, but unlike humans does not have the capacity to reflect on the ethical quality of what it does. Some philosophers therefore propose to build-in a capacity for ethical reasoning, [13,20], whereas other philosophers deny that this is possible or think it is insufficient for dealing with complex ethical issues in healthcare. On the other hand, maybe the robot's lack of moral agency is not a problem as long as humans are involved and included in the process. Again the issues of autonomy and role are raised.

- **Responsibility**. This issue raises again the question regarding the autonomy and role of the robot and the human and, more