Combining guidance ethics and the capability approach to inform the development of an AI-driven Clinical Decision Support aimed at promoting shared decision making on cardiovascular disease risks

André Krom[1], Myrthe van de Pavert[1], María Villalobos-Quesada[2]

[1] Department of Medical Ethics and Health Law, Leiden University Medical Center, the Netherlands
[2] National eHealth Living Lab, Public Health and Primary Care Department, Leiden University Medical Center, the Netherlands

Introduction: Clinical Decision Support Systems (CDSSs) assist healthcare professionals in clinical decision making. CDSSs can also be used to support shared decision making between doctors and patients. By integrating artificial intelligence in CDSSs (AI-CDSS), researchers aim to further strengthen shared decision making by enabling a personalised approach that considers the distinctive characteristics of each individual patient. Although AI-driven systems are potentially powerful tools in this regard, their adoption in clinical practice remains limited. The explanation for this is partly rooted in ethical considerations. Consistent with EU policy that AI must be responsible throughout its entire lifecycle, this paper presents the results of an ethical analysis of a specific use case: the early development of "The Clinical DECIsion support system for carDiovascular risk managEment in primary care" (AI-CDSS-DECIDE).

Methods: Data was generated using two Guidance Ethics workshops. Guidance Ethics (GE) is a process-oriented approach in which stakeholders systematically discuss ethical aspects of the possible use of a concrete technology in a concrete context (mapping actors, effects, values, and options for action). Workshops were recorded and transcribed verbatim, and researchers annotated the participants' perspectives in public during each workshop. Participants included designers, general practitioners, patients and patient representatives, decision makers, ethicists and AI and legal experts. For the second workshop extra attention was paid to being as inclusive as possible for participants who experience challenges with reading, writing and counting. An analysis of the transcripts was carried out using Atlas.ti, employing Martha Nussbaum's account of the Capability Approach (CA). CA was not explicitly mentioned during the workshops.

Results: In discussing potential effects of AI-CDSS-DECIDE, participants mapped direct and indirect effects on doctors, patients, their relationship and shared decision-making, as well as effects on the level of the healthcare system. In doing so, participants implicitly covered 7 out of 10 "basic capabilities" as identified by Nussbaum. Multiple personal and social conversion factors were implicitly discussed by participants as well. While not all workshop results could be described in terms of capabilities, functionings or conversion factors, interpreting the workshop results through the lens of the capability approach allows for a rich description not just of how and in which sense AI-CDSS-DECIDE could impact people's freedom to achieve wellbeing.
It also helps highlight relevant factors to take into account in shaping shared decision making regarding cardiovascular risk management, and the potential role that AI-CDSS-DECIDE can have in this. For instance, some potential effects, e.g. stigmatisation are effects in themselves, but can also have further effects – in this case as a negative social conversion factor resulting in chilling effects and feelings of shame, hindering people's freedom to achieve wellbeing. Reconceptualising the GE workshop results in term of the CA allows for fleshing out intrinsic and instrumental reasons to prevent stigmatisation, and to constructively examine possibilities to do so while also harnessing potential benefits of AI-CDSS-DECIDE. While the expected impact of DECIDE AI-CDSS on shared decision-making was considered mostly negative by participants, reconceptualising the GE workshop results in terms of the CA helps to put some of the negative effects identified in perspective. For instance, some effects considered to be negative ('If my GP calls me indicating that I have a higher risk of getting a heart attack or a stroke, my first response is panicking and thinking that I will die') could have positive effects down the line ('I will do everything that I did not come around to doing before'); indicating that it might not be necessary to prevent every and all negative effects (because some negative effects might at the same time serve as a positive conversion factor).

Discussion and conclusion:
Combining guidance ethics (GE) workshops with the capability approach (CA) seems promising for ethically evaluating AI-CDSSs. However, considering that not all workshop results could be interpreted in terms of the CA, the question arises as to whether it would be desirable to organise GE workshops in which the CA is incorporated beforehand.