Epistemic virtues and data-driven dreams: On sameness and difference in the epistemic cultures of data science and psychiatry

Marthe Stevens*, Rik Wehrens, Antoinette de Bont

Department of Health Care Governance, Erasmus School of Health Policy & Management, P.O. Box 1738, 3000, DR, Rotterdam, the Netherlands

ARTICLE INFO

ABSTRACT

Data science and psychiatry have diverse epistemic cultures that come together in data-driven initiatives (e.g., big data, machine learning). The literature on these initiatives seems to either downplay or overemphasize epistemic differences between the fields. In this paper, we study the convergence and divergence of the epistemic cultures of data science and psychiatry. This approach is more likely to capture where and how the cultures differ and gives insights into how practitioners from both fields find ways to work together despite their differences. We introduce the notions of "epistemic virtues" to focus on epistemic differences ethnographically, and "trading zones" to concentrate on how differences are negotiated. This leads us to the following research question: how are epistemic differences negotiated by data science and psychiatry practitioners in a hospital-based data-driven initiative? Our results are based on an ethnographic study in which we observed a Dutch psychiatric hospital department developing prediction models of patient outcomes based on machine learning techniques (September 2017 – February 2018). Many epistemic virtues needed to be negotiated, such as completeness or selectivity in data inclusion. These differences were traded locally and temporarily, stimulated by shared epistemic virtues (such as a systematic approach), boundary objects and socialization processes. Trading became difficult when virtues were too diverse, differences were enlarged by storytelling and parties did not have the time or capacity to learn about the other. In the discussion, we argue that our combined theoretical framework offers a fresh way to study how cooperation between diverse practitioners goes and where it can be improved. We make a call for bringing epistemic differences into the open as this makes a grounded discussion possible about the added value of data-driven initiatives and the role they can play in healthcare.

1. Introduction

Data-driven initiatives aim to analyse large volumes of data from varied sources to improve healthcare delivery by predicting future health risks and treatment responses (Mayer-Schönberger and Cukier, 2014; Murdoch and Detsky, 2013). Examples of such initiatives are machine learning applications that analyse medical images (Sample, 2018) or predict readmissions on intensive care units (PacMed, 2019) and artificial intelligence systems that assist in the diagnosis of cancer (Somsashekhar et al., 2017).

Such initiatives often come surrounded by promissory discourses that emphasize how large-scale or innovative data analysis will result in valid information to be used to enhance healthcare provision. These discourses result in a strong positive rhetoric about data-driven initiatives and as such drives their implementation. However, more cautionary discourses frame these initiatives in more critical terms and see the new initiatives as only a possible (or even a risk) to current

epistemic practices in the healthcare field (Stevens et al., 2018).

As fields originating in different disciplines, medical and data scientist communities have diverse epistemic practices, which come together in data-driven initiatives. Epistemic practices guide how members of a field propose, communicate, evaluate and legitimize knowledge. These epistemic practices are part of particular epistemic cultures (Knorr Cetina, 1981) that can be described as sets of specific norms, values, beliefs and traditions, that are "bonded through affinity, necessity and historical incidence" (Knorr Cetina, 1999:1). This means that epistemic cultures are known for specific activities for reasoning and establishing evidence, thereby determining what and how we know in communities.

The literature on data-driven techniques seems to either downplay or overemphasize the epistemic differences between data science and medical fields. On the one hand, data science scholars and some medical professionals easily dismiss their differences (e.g. Amato et al., 2013; Mayer-Schönberger and Cukier, 2014; Murdoch and Detsky,

* Corresponding author.
  E-mail addresses: stevens@eshpm.eur.nl (M. Stevens), wehrens@eshpm.eur.nl (R. Wehrens), debont@eshpm.eur.nl (A. de Bont).

2013). They imply that it is only a matter of time before the methods and standards of data science become part of established medical practice, thus suggesting an ultimate trajectory of convergence that downplays distinctions between their epistemic practices.

On the other hand, scholars are critical of the evidentiary claims of data scientists and primarily stress the epistemic differences. Some medical professionals explain in detail where and how the methods of both communities differ by theorizing on the specific role of data-driven approaches. They contrast grand concepts like causality and correlations and argue that data science can, for example, only be used to generate hypotheses and to explore valuable research directions (e.g. Khoury and Ioannidis, 2014). They argue for the specific use of data-driven approaches by highlighting the distinctions in terms of methods and evidentiary standards.

The aim of this paper is to study the concrete negotiation of differences in the epistemic cultures of data science and medicine without assuming either absolute incommensurability or deterministic complacency. Earlier studies about technological innovations have shown us that comparing (downplaying or emphasizing) the distinctions with the status quo is not helpful (Janssen, 2016; Smits, 2002; Van Lente, 2012) and we argue that it is more interesting to study where and how the epistemic cultures of data science and medicine overlap and differ. Such an approach is more likely to capture all the nuances, efforts and workarounds in practice. In addition, a more detailed empirical understanding of the negotiations in practice can also provide useful insights into how practitioners from diverse fields find ways to work with each other despite epistemic differences. This has broader implications for interdisciplinary knowledge practices that are increasingly more common and more expected (e.g. large European projects) (Rathenau, 2018).

We conducted a case study of a data-driven initiative in the context of psychiatry. Psychiatry is a particularly interesting case to study the introduction of data-driven initiatives as this field is characterized by considerable uncertainty relating to disease ontology and treatment effects. Moreover, it relies primarily on narratives of patients and qualitative questionnaires to make sense of patients' conditions and guide treatment decisions. This means that uncertainty related to data collection and use, that is also part of other fields of medicine, is amplified in the field of psychiatry.

The paper is based on an ethnographic, empirical study of the development and implementation of a hospital-based data-driven initiative in the Netherlands. Within this initiative, data scientists were brought into a psychiatric hospital department to develop prediction models of patient outcomes based on machine learning techniques. We approach this initiative through the sensitizing concepts of "epistemic virtues" (Daston and Galison, 2007) and "trading zone" (Galison, 1997). The notion of epistemic virtues enables us to ethnographically focus on the differences in epistemic cultures, while the concept of trading zones allows us to zoom in on how the differences are negotiated. This leads to the following research question: *how are epistemic differences negotiated by data scientists and psychiatrists in a hospital-based data-driven initiative?*

In the next section, we sketch the theoretical background and elaborate on epistemic differences by introducing the concepts of epistemic virtues and trading zones. We explain how this combination helps analyse the practices in the initiative. This is followed by the case description and methodology. In the results, we introduce the data-driven initiative and present two cases from our fieldwork that illustrate the process of trading epistemic differences. The discussion concentrates on the role of epistemic virtues and how they play a role in interdisciplinary cooperation.

2. Theoretic framework

2.1. Epistemic differences

A central theme within science and technology studies (STS) has been critiques and reflections on the idea of "epistemic unity of the sciences" (Galison and Stump, 1995). In this tradition, Knorr Cetina (1981; 1999) developed her work on "epistemic cultures". She is known for her ethnographic comparison between experimental high-energy physics and molecular biology. She shows that knowledge is created in diverse scientific cultures and this results in a wide array of scientific practices and preferences that all coexist. Building on this body of work, we conceptualize data science and psychiatry as two different, but not necessarily incommensurable cultures of knowledge with their own epistemic practices.

This conceptualization of data science and psychiatry as diverse epistemic cultures is firmly grounded in literature. Scholars investigating data science have, for instance, pointed out how the field is characterized by an epistemic tradition with a specific history that is different from the epistemic culture described in literature about psychiatry. It is known that data science is a relatively new field (Baru, 2019) that grew with the capacities to gather and analyse large amounts of data and the increasing "datafication" of aspects of our life (Mayer-Schönberger and Cukier, 2014). It originated from computer science, statistics and mathematics, but more explicitly engages with "real" data under "real-world" constraints (Baru, 2019). A strong positive rhetoric and commercial successes led to the increased realization that real data are a valuable commodity and new methods, infrastructures, technologies and skills were being developed to handle these data (Leonelli, 2014). The field is described as relatively a-theoretical, as it is a general approach that can be applied to analyse data drawn from a wide variety of fields and domains (Ribes, 2019).

Epistemically, the field of data science assumes that there is a (be it, complex and multifaceted) real world or reality that can be better or more completely captured by real-world data (Stevens et al., 2018; Leonelli, 2014; Ribes, 2019). Thus, the field assumes that there is a strong connection between the real-world and data. This sentiment leads to the suggestion that we can know more and act better from the careful grouping and analysis of real-world data and extrapolating them into the future. This commitment to these data means that this field is (in principle) respectful of any statistical relationship between two data values and accepts that some relations might never be understandable or explainable (Leonelli, 2014).

The field of psychiatry has a longer history than the field of data science. Since the twentieth century, the field of psychiatry has undergone numerous and by no means monolithic transformations (Rüppel and Voigt, 2019). During this time, psychiatry has been influenced by the evidence-based medicine (EBM) movement that resulted in a preoccupation with randomized controlled trials (RCTs) and experimentation (Timmermans and Berg, 2003; Rüppel and Voigt, 2019) and also by the "molecularization of the "medical gaze" (Rose, 2007) that led to studies into genetics and biomarkers that aimed to localize diseases within the body (Rüppel and Voigt, 2019). However, the EBM-movement had more impact in other fields, as measurements could be obtained under more controlled conditions and with more exactitude and accuracy. Therefore, psychiatry can still be characterized by considerable uncertainty relating to disease ontology, treatment effects and strong reliance on patient narratives.

Epistemically, the field of psychiatry assumes that diverse genetical, neurobiological, environmental, biographical factors and the interactions between them lead to the development of psychiatric diseases in some people (Rüppel and Voigt, 2019). Thus, the field assumes that there are (strong) causal connections, but they are complicated and often unknown. This sentiment results in diverse kinds of research that aim to capture relationships between events and outcomes. These studies and their evidence are consequently ranked and evaluated based on their quality and conventions (Vidal and Ortega, 2017). The preference for specific research methods means that the field is critical about scientific methods (e.g. large observational studies and RCTs are preferred above individual case studies). This contributes to gaps in evidence that need to be bridged by clinical interpretation in practice (Rüppel and Voigt, 2019).

This literature suggests that both data science and psychiatry assume a strong and complex connection between events in the world and data that are gathered about them. These data can be analysed and are, in principle, a sound basis for decision-making. However, there are differences in the sort of data and methods that are used and how they are valued.

2.2. Epistemic virtues

In order to understand where and how the epistemic cultures of data science and psychiatry differ and to study how these differences are concretely negotiated in data-driven initiatives, we build on the concept of epistemic virtues. Research in philosophy of science and STS have shown that epistemic norms and values play a role in scientific communities and guide what is perceived to be normal and acceptable (Kuhn, 1962; Latour, 1987).

Epistemic virtues can be conceptualized as epistemic norms and values that are internalized and acted upon by data scientists and psychiatrists (Daston and Galison, 2007:40–41) and are one of the ways of studying epistemic norms and values. Both data science and psychiatry have internalized norms about, for instance, certainty, representativeness and objectivity. These can be understood as virtues that are acted upon in specific ways, for example, in the judgements that medical practitioners pass on one another's work. They do not use the word "virtue" or "vice", but their praise or blame often relates to qualities of work that they consider "good" or "bad" (Paul, 2011:7). The concept of epistemic virtues thus helps us to look ethnographically at precisely those moments when specific judgements are made by data scientists and psychiatry practitioners. Epistemic virtues enable us to interpret such moments as not merely a methodological discussion, but as examples that signal underlying differences in epistemic cultures.

In contrast to approaches that tend to list epistemic virtues considered relevant for knowledge gathering (e.g. Marcum, 2017; Pigliucci, 2017), our use of epistemic virtues stresses the dynamic and social character of epistemic virtues (and hence also epistemic cultures). This understanding of the notion emphasizes that: (1) epistemic virtues are actively constructed and continuously re-evaluated and, (2) epistemic virtues are situated, which means that people, often together, always give meaning to epistemic virtues in specific circumstances.

We are inspired by the work of Daston and Galison (2007) around the notion of "objectivity". They showed how the meaning of the epistemic virtue "objectivity" was (re)constructed throughout the years and showed how objectivity was understood differently by scientists in the eighteenth, nineteenth and twentieth centuries. Their approach inspired us to look at the multiple interpretations that are given to epistemic virtues and what this says about sameness and differences of two epistemic cultures. As a useful addition to the work of Daston and Galison (2007), we suggest that the notion of epistemic virtues is not only useful to study longitudinal changes in scientific development, but can also be fruitfully applied to study the in situ negotiation of differences in epistemic cultures (Knorr Cetina, 1981).

2.3. Trading zones

While the notion of epistemic virtues enables us to ethnographically focus on the differences in epistemic cultures through concrete moments of judgment, the concept of trading zones allows us to zoom in on how these differences are negotiated. We frame the data-driven initiative as a trading zone (Galison, 1997). Galison (1997) also challenged the "epistemic unity of the sciences" in his work about trading zones as he was intrigued by the extraordinary variety and disunity of scientific languages and practices (Galison and Stump, 1995). He analysed how distinct communities in physics – such as theorists, experimentalists and engineers - create in-between vocabularies that facilitate communication and alignment of activities (Galison, 1997). These intermediating languages can range from simple (interlanguages) to complex ("pidgin") and eventually a shared language can emerge ("creole") (Collins et al., 2007; Galison, 1997). These languages

make it possible to interact and exchange goods despite differences and without homogenizing the inherent diversity in their communities (Galison, 1997).

Galison's notion of trading zone is used as a tool to analyse a wide variety of interactions between different communities, ranging from cross-boundary interactions between surgeons and engineers (Baird and Cohen, 1999), ways in which team members with different backgrounds cooperate in a marketing firm (Kellogg et al., 2006) to communication between NASA engineers and their subcontractors (Vaughan, 1999). This body of (ethnographic) work continuously stresses that the different communities do not meet each other "with gaping incomprehension" (Harmen and Galison, 2008:568), but that members of the diverse communities coordinate their actions temporarily and locally. They navigate their differences in language and culture only as needed (Kellogg et al., 2006; Galison, 1997) and this is exactly what gives science its strength and coherence and underlies the experience of scientific continuity (Galison and Stump, 1995).

The trading zone literature pays attention to the linguistic and material components that help cross-boundary interactions. Galison emphasized the importance of language right from the start. Later, Collins et al. (2007) expanded on the linguistic understanding of trading zones with their work on "interactional expertise". This concept highlights two things. First, that not only communities but also independent third parties can gain interactional expertise in talking to both communities in some approximation of their language (Collins et al., 2007; Collins and Evans, 2008). Second and more importantly, interactional expertise emphasizes the social processes and tacit knowledge that are passed on through language and that are important to facilitate exchange (e.g. Epstein, 1996). This work highlights that socialization is important for learning another culture because through day-to-day immersion, people learn and understand the rules that cannot be written down (Collins and Evans, 2008; Wehrens, 2015).

Besides the linguistic components, there has been a focus on material, "boundary" (Star and Griesemer, 1989) objects that facilitate exchanges largely in the absence of linguistic interactions. The objects facilitate trading because they are "plastic enough to adapt to local needs and constraints of the several parties employing them, yet robust enough to maintain a common identity across sites" (Star and Griesemer, 1989:411). The objects can mean something distinctively different to both parties and help interactions because they do not vitiate the communities' separate projects (Collins et al., 2007).

In this paper, we focus on the negotiation of epistemic differences and while language and objects can help to facilitate exchange, much is still unknown about the role of epistemic virtues in trading zones. Galison (1997) remarks that communities can have diverse norms, values and virtues. According to Galison (1997:401, 807), the communities persistently try to incorporate the virtues of the other. However, a more in-depth exploration of the role of epistemic virtues in trading zones still has to be developed. We aim to contribute to this development by paying attention to precisely those moments in the initiative when specific judgements are made by data scientists and psychiatry practitioners. We argue that the focus on epistemic virtues as concrete moments of judgement can be a useful additional dimension in the empirical study of trading zones, next to linguistic and material dimensions developed by other authors.

3. Methods

3.1. Case description

This study focuses on a data-driven initiative by a psychiatric department in one of the largest Dutch University Hospitals. Combining outpatient and inpatient treatment, this department contains four specialist units that specialize in affective, psychotic and developmental disorders, acute and long-term care. The department treats approximately 2000 unique patients annually. There is a strong focus on research in the department as it is linked to various research groups, and

Table 1
Overview of key practitioners of both communities.

Practice	Practitioners	Main involvement	Data collection
Data science	Internal data scientist A	Team leader, all research projects	Interview, observations (17 days)
	Internal data scientist B	Antipsychotic project	Interview, observation (18 days)
	Internal data scientist C	Antipsychotic project	Interview
	Internal data scientist D	Antidepressant project	Interview, observation (16 days)
	External data scientist E	Antipsychotic project	Interview, observation (6 days)
	External data scientist F	Antipsychotic project	Interview, observation (7 days)
	External data scientist G	Antidepressant project	Interview, observation (11 days)
	PhD, data scientist H	Various projects	Interview, 12 days present in observations
Psychiatry	Psychiatrist I	Professor, medical head of department, initiator of the data-driven team, various research projects	Interview, observation (4 days)
	Psychiatrist J	Department head, studied schizophrenia, antipsychotic project	Observation (2 days)
	Psychiatrist K	Studied social cognition in schizophrenia, antidepressant and antipsychotic project	Observation (3 days)
	Psychiatrist L	Studied gene-environment interaction and antipsychotic project	Observation (3 days)
	Psychiatrist M	Studied developmental disorders in children. Advisor to the data-driven team	Interview
	Psychiatrist N	Studied genetic predispositions in relation to mental illnesses, antidepressant project	Interview, observation (3 days)
	Researcher O	Assistant professor, expertise in neuroimaging. Advisor to the data-driven team	Interview
	Researcher P	Assistant professor, expertise in genetic epidemiology, antipsychotic project	Observation (2 days)
	Nurse Q	Involved in antidepressant project	Interview, observation (1 day)
	Nurse R	Advisor of the data-driven team	Interview

many psychiatric practitioners are also involved in research. Studies focus, for example, on genetic predispositions, brain morphology, and risk factors that play a role in the development of psychiatric disorders. The department is well-established as a research centre, both in the Netherlands and worldwide.

The data-driven initiative was initiated by the medical head of the department in 2015 as she set out to experiment with data-driven techniques to improve treatment for psychiatric patients, in addition to research initiatives already being conducted in the hospital. The department attracted data scientists who had worked for the major companies or were involved in data science start-ups. The initiative secured the help of an IT company and acquired funding from the Dutch government and other organizations. With these funds, the data scientists could start pilots, build structures that automatically supplied data and start working toward making care improvements for patients.

When we followed the initiative, the core of the data-driven team comprised four data scientists, that were supported by multiple students, six external data scientists and one data engineer. Students (e.g., medical, psychology) helped the data scientists with small projects. The six external data scientists were hired to work on short-term subprojects with specific goals in mind (they followed the funding). The data engineer worked for the university that was linked to the hospital and provided technical support.

At the time, the team worked on several data-driven subprojects. Besides presenting straightforward business analytics (e.g. how often is a certain medicine prescribed?), the team worked on developing decision-support systems that would aid psychiatrists by making predictions based on machine learning algorithms and experimented with the implementation of the machine learning models. Two major subprojects tried to predict which antipsychotic or antidepressant would be the most effective while producing the least side effects for certain patients. These projects were relevant because they would (1) avoid the current trial and error method, where psychiatrists must try several drugs before finding the right one; (2) reduce hospital stay; (3) be expandable to other psychiatric medication in the future.

3.2. Data collection

Our empirical analysis builds on a combination of qualitative methods. First, MS observed the data-driven team between September 2017 and February 2018 (approximately 200 h). After an initial meeting with the medical head of the department and the leader of the data-driven team in which we explained our research focus, we could meet and shadow the data-driven team. MS shadowed the data scientists while they were doing their analyses and followed them around to

meetings in the hospital, taking field notes to capture the setting and interactions between data scientists, medical staff and researchers (Oldenhof, 2015). MS introduced herself and the research to all involved and made sure ongoing informed consent was obtained. Informal conversations helped build rapport and provided insights into the pertinent issues and tensions. The field notes were expanded as soon as possible and discussed with RW and AdB.

Second, MS conducted seventeen semi-structured interviews with nineteen people (directly and indirectly involved with the data-driven initiative) to explore particular topics in more depth: nine data scientists, three psychiatrists, two nurses, one data engineer, one data manager, one consultant, one hospital manager and one medical researcher. The interviews posed open questions about the initiative and the collaboration, such as: how do you go on collaborate with others? All interviews lasted 40–60 min and were recorded and transcribed ad verbatim. We asked permission for the interviews, the use of quotes and anonymized the material.

Finally, MS conducted an analysis of online and offline documentation about the initiative, such as presentations and newsletters. This resulted in more than 400 pages related to the initiative. An ethical waiver was obtained for the study.

3.3. Data analysis

We began our analysis with open coding (Mortelmans, 2007). We performed general readings and highlighted notable passages in the data. For example, we marked passages in which practitioners from one community expressed that they did not understand the other community. MS conducted most of the analysis, but researcher triangulation ensured that key themes emerging from the analysis were discussed and refined (Mortelmans, 2007). Through several iterations of the analytical process, we focused on three key categories: (1) epistemic claims and ideas, (2) the negotiations and discussions between data-driven and psychiatry practitioners, and (3) the coordination of the meanings and ideas about knowledge production.

Although analysis began with open coding, the total process employed an abductive approach involving an iterative to-and-from between analytical themes and relevant theoretical concepts (Timmermans and Tavory, 2012). After identifying the literature on trading zones and epistemic virtues as crucially relevant for the analysis of the empirical material, we developed the analysis further through posing such questions; (1) what is traded here, by whom and where? (2) Which epistemic virtues are in play? (3) What does this say about the respective epistemic cultures?

In the end, our analysis focused on eighteen key practitioners involved in the initiative. Table 1 provides an overview of the practitioners per community, their role and data collected about them. The remaining people

under observation had a minor role in the trading process. Their interviews helped to gain more information about the structure of the trading zone and details of the various projects. Due to time constraints and high workloads, not all psychiatric practitioners could be interviewed.

We incorporated member checks in various phases. During data collection, we kept in touch with two (key) respondents to clarify unclear situations. After the data collection, we presented our initial results and a first draft of the paper to gain feedback on our analysis. The respondents recognized their work in the analysis, and pointed out minor misunderstandings (e.g. datasets were updated on a weekly basis instead of daily basis). We presented the final draft of the paper to our respondents to gain their permission to use quotes and asked them to check our use of this material (Mortelmans, 2007).

4. Results

In the next section, we will present the hospital-based data-driven initiative and the data science and psychiatry practitioners involved. Afterwards, we analyse the process of trading epistemic differences, illustrated by two examples derived from data science subprojects developed in the initiative.

4.1. The trading zone

We learned that not all the psychiatry practitioners greeted the data scientists with enthusiasm, when they arrived at the hospital. There were diverse criticisms. Some psychiatrists were sceptical because of previous costly, time-intensive ICT projects had limited added value. Others said that the rapid nature of developments in data analysis would quickly outdate any initiative. They argued it would be better to wait and see how techniques developed in other contexts. Also, some questioned the methods used. Psychiatrist M explains:

"When they first presented the ideas [on the initiative], I really had to hear them a few times and think about them. That's how it works for me! (…). I've learned to be very suspicious of connections that you find in data. The more you look for certain connections, the greater the chance that you will find one, but it could well be complete nonsense." (Interview psychiatrist M)

The data scientists decided to focus their attention on psychiatry practitioners who were enthusiastic about the initiative and willing to cooperate. The psychiatry practitioners who participated in the data-driven initiative hoped to find new sorts of information that could help them to improve the treatment of patients. Several psychiatrists described the unpredictability of psychiatric diseases. They mentioned, patients with long-lasting depression who unexpectedly recovered, patients who benefited from strange combinations of medicines or patients who developed severe, unanticipated side effects. Psychiatrists relied on a combination of scientific evidence and their professional experience to treat patients but realized that their experience did not always provide clear evidence and contained biases. As psychiatrist I explained:

"We try to treat patients to the best of our ability, but often we would like to have more information. There are guidelines for extreme situations, but there is a large gray area. Those guidelines are fine, we don't need to change them, but we do need more guidance in this gray area. We don't know much; there is simply not a lot of evidence. We need some sort of … intermediate evidence." (Interview psychiatrist I)

The data scientists kept in touch with the medical practitioners through face-to-face meetings, via e-mail and by accompanying them on the daily rounds. We observed practitioners in both communities discussing the data that could be included in a prediction model, validating preliminary machine learning outcomes and brainstorming on verifying the decision-support models. We highlight two cases: the first case is illustrative of the majority of the negotiations we observed. The second case because negotiations were most fierce and challenging.

4.2. Negotiations during the development of a prediction model for antidepressants

The first case is of a typical meeting between data scientists and psychiatrists. The meeting is part of the trajectory in which the data scientists are trying to develop a prediction model that will help determine which antidepressant is best for certain groups of patients. During this meeting, data scientists (A, D and G) and psychiatrist N are sitting in N's small office. The data scientists have extensively studied prescription data and guidelines and have had similar meetings with other psychiatry practitioners. Now, they want to validate their findings and come to a decision about how to proceed.

Case 1: How to define a successful treatment?

The meeting starts with a discussion about prescription behaviour and the possible side effects of antidepressants. Data scientist G explains that the data scientists are trying to distinguish between "successful" and "unsuccessful" medical treatment as the first step in developing their prediction model. After consulting with various psychiatrists and studying the guidelines, he is considering the following conditions. First, a drug must have been prescribed for a minimum amount of time, so that it has enough time to be effective. Second, the Hamilton depression score [a measure of the severity of depression] must be at least 50% lower than on the patient's admission date or should be equal to 8 [indicating sufficient improvement or a "normal" value]. He is thinking of including Beck scores and the Functional Disability Inventory [both aimed at measuring depression severity]. The only problem is that they are embedded in text fields in the patient files, which makes them hard to use. Data scientist D adds that they have heard from other psychiatrists that MADRS [another depression scale] might be relevant. However, they noticed in the data that medical professionals in this hospital do not use the MADRS scale.

Data scientist G argues that the Hamilton score seems most important. Psychiatrist N nods in agreement, but sighs: "What a shame that it's so hard to analyse free text fields in patient files." Data scientist G explains that they are unsure about the "50% lower" or "equal to 8" criteria as these seem quite random threshold levels. Psychiatrist N disagrees because, according to him, most medication studies use the same threshold levels. He gives some context: "A Hamilton score can go up to around 25 points when someone is severely depressed. You should be more worried about measurement variation because medical students fill out the questionnaire with the patients during their internships and they don't always take it seriously." This is new to the data scientists, but, as data scientist A remarks; "We do not really have another measure that we can use, so this is our best option." Psychiatrist N agrees.

Data scientist G shows a graph of the current distribution of Hamilton scores for patients treated with antidepressants [see Fig. 1 for a similar but fictional graph]. The graph reveals that only a small portion of patients lands in the "lower or equal to a score of 8" (area 1), meaning that they have had "successful treatment" according to the current criteria. A more substantial group of patients have had successful treatment if "50% improvement" (area 2) is also considered. A few patients show a partial response (area 3), and even fewer show no or a negative response (area 4). The data scientists suggest also including patients with partial response (area 2 and 3) in their definition of "successful", as their Hamilton scores improved after medical treatment.

Psychiatrist N looks at the graph more closely and remarks that the graph is different than what he would have expected, based on the scientific literature. He asks the data scientists what sorts of data they used to make the graph. He goes to his computer, intending to look up a scientific study, when suddenly his phone rings, interrupting the meeting. Something is happening to one of his patients. It seems quite serious because psychiatrist N excuses himself and, heading off to the ward, adds, "We must continue this conversation next week." (Observation notes).
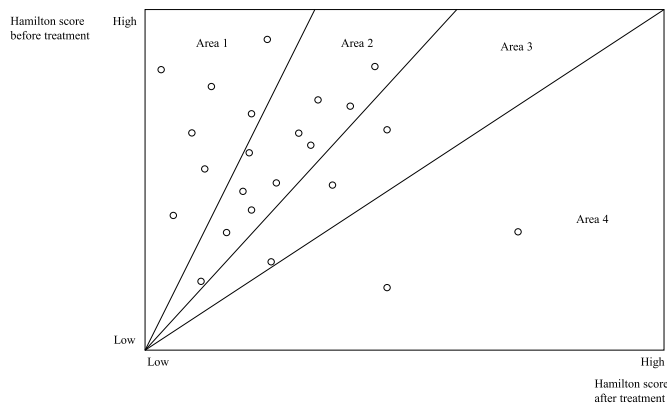
Fig. 1. Distribution of Hamilton scores for patients treated with antidepressants. Each dot represents an individual patient.

This first case shows how the data scientists and psychiatrist negotiate on the data to include and adequate threshold levels. During the meeting, they were able to navigate some of their differences and make the decision to base the prediction model solely on the Hamilton data. At one point, the data scientists were afraid to lose their last data source (Hamilton data) when psychiatrist N started to talk about measurement variation. At that moment, both parties decided pragmatically to use the Hamilton data as it was "the best option".

The trading process took place during formal meetings and it helped that the data scientists learned to speak in the vocabulary of questionnaires and depression scales that is understandable for the psychiatrist. Simultaneously, psychiatrist N confirmed the struggles of the data scientists and contributed to the negotiations by stating "what a shame that it's so hard to analyse free text fields", thereby showing that he understood the importance of these data for the data scientists. The trading process was also stimulated by the *systematic approach* of the data scientists. This virtue was recognized by psychiatrist N. He understood the importance of determining outcome measures, as it was also an important part of the scientific research he conducted in the past. The graph also helped to visualize the threshold dilemma and thereby contributed to the cross-boundary interactions.

There were also epistemic virtues that needed to be negotiated. The data scientists looked for *completeness* in the data. They wanted to include as much data about depressions as possible; several measurements, qualitative data and data of patients that partially respond to the treatment. At the same time, psychiatrist N was more *selective* about the data to use and threshold levels to select. N referred to scientific studies during the meeting. With his mentioning of scientific literature, he brought in authority and the rich scientific history of psychiatry. He stressed the importance of being selective in the kinds of data to include and thresholds to uphold while simultaneously highlighting the embeddedness of psychiatry in practice.

4.3. Negotiations during the implementation of a prediction model for antipsychotics

The second case highlights a trading process about the importance of statistical significance between data science and psychiatry practitioners. The discussion became most evident during the antipsychotics subproject. Within this project, both parties had to negotiate the standards and measures to uphold for determining the performance of the prediction model and the conditions under which the model could be implemented in healthcare practice.

Case 2: When to introduce a machine learning model in daily healthcare routines?

The psychiatry practitioners needed certainty that the models would significantly improve current healthcare practices. They argued that, while not perfect, statistical significance would be a necessary step to test the performance of the models before introducing them in daily healthcare practices, since it would help determine the reliability and ensure that the success of the model was not based on chance. They argued that it was necessary to set up a RCT-inspired approach, in which the performance of the model would be tested against care-as-usual.

The data scientists did not completely understand the focus on statistical significance. Some of the data scientists previously worked in the commercial sector, where they never had to work with statistical significance. They searched for findings with the biggest impact and tended to look at prediction accuracy instead of statistical significance. Data scientist G explained: *"If you started to talk about significance in such a [commercial] setting, everybody would look at you and say: "what are you talking about?" (Interview data scientist G).* The data scientists explained that they use their own methods to validate their findings and to see if their model is performing well:

Data scientist E: "Significance is just not that interesting with techniques like machine learning! Well, that might be a bit exaggerated, but …"

Data scientist C: "It is not the ONLY measure!"

Data scientist E: "We often use lots of data to generate our models, when significance is not necessarily the most important measure. There are other ways to validate that your model is doing well and in what circumstances." (Interview data scientists C and E).

The notion of statistical significance caused misunderstanding and miscommunication. At the beginning of the project, the data scientists had proudly presented some of their initial findings when a psychiatry practitioner asked if the results were statistically significant. The data scientists answered that such a statistical test was less important nowadays. It was more important to validate your findings with medical professionals. Psychiatric researcher O was critical about this approach:

"I don't agree with the approach. I mean, I think you still have to make sure that a result is statistically significant and not produced by chance, so I wonder about dismissing this statistical test as somewhat unimportant." (Interview psychiatric researcher O).

The presentation of the data scientists had ended with a fierce discussion and became a story that was retold in both communities.

During the end of our fieldwork, uncertainty remained about the conditions under which the data scientists could implement their machine learning models in practice. The psychiatry practitioners continued to hold the opinion that statistical significance was crucial for the evaluation of the prediction models. Some argued that data scientists should not even present non-statistically significant findings to them, as this could negatively influence their behaviour. Data scientists stated, by contrast, that *"it would be a pity not to use all the non-significant results. You waste a lot of information that could be used to our advantage"* (data scientist F).

The data scientists tried to obtain some leeway by organizing meetings with psychiatry practitioners that were interested in their methods. This allowed the data scientists to explain their alternative validation methods. Also, strategic meetings with more senior staff were organized to bring both parties together and make workable agreements. Simultaneously, the data scientists agreed with the RCT-approach to compare their models with care-as-usual as they understood the benefit of such a comparison.

This second case shows how data science and psychiatry practitioners negotiated about the implementation of the machine learning models in practice. There was a partial compromise as both parties agreed on the RCT-approach to compare the prediction model against the care-as-usual. The data scientists did see the value of this approach and understood the responsibilities of the medical practitioners. However, the negotiations were far more intense and developed differently than in the first case.

In this case, trading was for the large part done by backstage politics. The data scientists did set up meetings with interested medical practitioners to make them familiar with their epistemic culture and to teach them their "language". Besides, the strategic meetings with senior staff were organized. These persons were respected by both communities and could stimulate the negotiations. Also it helped that both the data science and psychiatry practitioners felt that data, models and outcomes had to be verified and recognized the importance of having *validation standards* as a virtue. The negotiation process that unfolded was about the sort of validation process to use.

There were epistemic virtues that needed to be negotiated and this was not straightforward. The psychiatry practitioners, while enthusiastic about the prediction models, were reluctant to introduce these models into their care routines. They needed *certainty* about the functioning of the model and the methodologies used. This was not necessarily unwillingness from their part, as it also ties closely to the various epistemic responsibilities and risks that they face (e.g. treatment responsibility and malpractice claims). The data scientists could be more *flexible* in the outcomes and evidence they selected as they do not face these responsibilities and risks. The flexibility was also visible in the statement of data scientist E when he talked about validating that "the model is doing well in certain circumstances". This shows that, for the data scientists, the necessary degree of flexibility or certainty was dependent on the impact of the analysis. This would mean that a data-driven model for deciding on deep brain stimulation needed more assurances than a model used for scheduling staff. This more dynamic use of information meant that data scientists often accepted more flexibility in their methods and stressed that they might have to revise their analyses after a few years or based on new data.

The trading process, in this case, was more difficult than in the first case. It was not helpful that both communities retold stories (such as about the discussion on statistical significance) within their own communities as it emphasized the differences between the communities. The negotiations were already difficult, especially because the virtues (flexibility versus certainty) were so different and complex statistical theories clouded the discussions.

The field of psychiatry is used to be working with causal inference and research that looks for causation. This type of research aims to give answers to questions such as: "how much more effective is antipsychotic A compared to antipsychotic B for (a group of) patients?". Medical practitioners can use this information to *change* the treatment practice for their patients. The practice of prediction modelling of the data scientists is based on another statistical theory. This type of research tries to predict what the chance is that someone improves when prescribed a particular type of antipsychotic. As such, it allows medical practitioners to be *reactive* to situations. It, for example, presents the chance that a side effect will occur in a patient and this enables psychiatrists to anticipate on the side effects. The discussions about significance illustrate that it is complicated to understand and act upon these different sorts of information. Causality thinking forms such a considerable aspect of the culture of psychiatry and is so embodied by all practitioners, that might be difficult to grasp alternative approaches and vice versa.

5. Discussion

In this paper, we study the convergence and divergence of the epistemic cultures of psychiatry and data science. We argue that such similarities or differences are not pre-given but depend on situated negotiations and actions. Through the notions of "trading zones" and "epistemic virtues", we have been able to analyse how differences between epistemic cultures are negotiated in situ by data science and psychiatry practitioners. In this discussion, we argue that our combined theoretical framework offers a fresh way to study how cooperation between diverse practitioners develops and where it can be improved. We make a call to bring epistemic differences into the open as it makes a grounded discussion about the added value of data-driven initiatives and the role they can play in healthcare, possible.

Our ethnographic approach showed that practitioners from both communities actively sought collaboration to improve care for psychiatric patients. Many epistemic differences needed to be negotiated. Situated judgments about the need for being complete or selective in the data to include and about the amount of flexibility or certainty that is permissible, revealed the role of specific epistemic virtues. We found that it is not simply the case that diverse communities persistently try to incorporate the epistemic virtues of the other (Galison, 1997:401, 807). The dynamics are more complex and far messier as epistemic virtues get their meaning in concrete practices.

Our study of the data-driven initiative shows how epistemic differences are traded locally and temporarily. Our work aligns with earlier work on trading zones in highlighting the importance of boundary objects (the graph in the first case), interactional expertise (data scientists learning the medical language in the first case) and socialization (data scientists becoming embedded in the culture of psychiatric measurements in the first case) in stimulating cross-boundary interactions (Galison and Stump, 1995; Galison, 1997; Knorr Cetina, 1999). Our research adds to this literature by highlighting two additional processes. First, our work shows the importance of validation of the other culture in helping to navigate differences. The remark of the psychiatrist in case 1 about how unfortunate it was that free text fields are so hard to analyse, showed that he learned enough about the culture of the data scientists, and by explicating these differences he showed respect for the other culture. Second, shared epistemic virtues, such as the systematic approach and the validation standards, helped to stimulate the negotiation processes and create space for negotiation, despite differences.

In both cases, the communities were able to navigate some of the differences in culture. However, the negotiations in the second case were more complicated, because epistemic virtues were too different (e.g. flexibility and certainty) and hidden within complex statistical theories. What also complicated the interactions, was the retelling of stories that emphasized the epistemic differences.

The study has two limitations. First, our distinction between the two subcultures might be considered too simplistic or artificial as there is never just one single data-driven culture nor only one medical culture (Harmen and Galison, 2008; Knorr Cetina, 1981). We argue that while there is variation, there is ample evidence that shows an overarching medical and data-driven epistemic culture as we described in our theoretical framework. For the purpose of this study, aimed at exploring the similarities and differences in epistemic culture, we drew a division based on the participants' positions in the hospital-based initiative, but different divisions could be made.

Second, while rich in detail and extended with other data collection methods, our ethnographic study of this data-driven initiative lasted six months. The relatively short timeframe made it impossible to observe, more longitudinal process changes, such as the development of shared language or more gradual changes in epistemic virtues. As such, more longitudinal ethnographic research projects, following data-driven initiatives over an extended period, would be a welcome addition.

We argue that it is important to bring epistemic differences into the open as this enables a grounded discussion about the added value of data-driven initiatives and the role they can play in healthcare. This study shows the added value of ethnographic, action-based research in making subtle differences in epistemic cultures visible in concrete data-

driven initiatives. Data-science initiatives are not impossible or simply possible in healthcare and epistemic virtues help to study where problems arise and how practitioners deal with them. This approach helps to make a more grounded discussion possible, about the type of data to include, thresholds to uphold and the necessary steps for the implementation of prediction models.

Such ethnographic studies can subsequently be used to organize moments of reflexivity for data-science and healthcare practitioners, for example, by starting initiatives and experiments that stimulate discussions. Think of organizing focus groups and data deliberation sessions, involving the relevant communities, such as professionals, patients, data scientists, researchers and policymakers, to explore the epistemic differences and similarities (Haan et al., 2018; Madsen and Munk, 2019; Moats and Seaver, 2019; Ziewitz, 2017). Epistemic cultures are not fixed, and they can change. The epistemic cultures of data science and psychiatry are in some respects relatively similar but also differ in myriad ways. It is highly relevant to study the differences and see how both cultures can learn from each other.

Author contribution

Marthe Stevens: conceputalizations, Investigation, Writing - original draft preparation, Project administration. Rik Wehrens: Conceptualizations, Writing - review & editing, Supervision. Antoinette de Bont: Conceptualizations, Writing - review & editing, Supervision.

Declaration of competing interest

None.

Acknowledgements

References

Amato, F., López, A., E M, Peña-Ménez, Vaňhara, P., Hampl, A., Havel, J., 2013. Artificial neural networks in medical diagnosis. J. Appl. Biomed. 11 (2), 47–58.

Baru, C., 2019. Data Science Needs a Clinical Program. Retrieved from [March 2020]. https://chaitanbaru.blogspot.com/2019/02/.

Baird, D., Cohen, M.S., 1999. Why trade? Perspectives on Science 7 (2), 231–254.

Collins, H., Evans, R., Gorman, M., 2007. Trading zones and interactional expertise. Studies in History and Philosophy of Science 38 (4), 657–666.

Collins, H., Evans, R., 2008. Rethinking Expertise. The University of Chicago Press.

Daston, L., Galison, P.L., 2007. Objectivity. Zone Books.

Epstein, S., 1996. Impure Science: Aids, Activism and the Politics of Knowledge. The University of California Press.

Galison, P.L., 1997. Image and Logic: A Material Culture of Microphysics. The University of Chicago Press.

Galison, P.L., Stump, D.J., 1995. The Disunity of Science: Boundaries, Contexts, and Power. Stanford University Press.

Haan, M.M., van Gurp, J.L.P., Naber, S.M., Groenewoud, A.S., 2018. Impact of moral case deliberation in healthcare settings: a literature review. BMC Med. Ethics 19 (85), 2–15.

Harmen, O., Galison, P.L., 2008. Epistemic virtues and leibnizian dreams: on the shifting boundaries between science, humanities and faith. Eur. Leg. Forum 13 (5), 551–575.

Janssen, M., 2016. Situated Novelty: A Study on Healthcare Innovation and its Governance. Erasmus University, Rotterdam.

Kellogg, K.C., Orlikowski, W.J., Yates, J., 2006. Life in the trading zone: structuring coordination across boundaries in postbureaucratic organizations. Organ. Sci. 17 (1), 22–44.

Khoury, M.J., Ioannidis, J.P.A., 2014. Big Data meets public health. Science 346 (6213), 1054–1055.

Knorr Cetina, K., 1981. The Manufacture of Knowledge: an Essay on the Constructivist and Contextual Nature of Science. Pergamon Press.

Knorr Cetina, K., 1999. Epistemic Cultures: How the Sciences Make Knowledge. Harvard University Press, Cambridge, Massachusetts, London.

Kuhn, T.S., 1962. The Structure of Scientific Revolutions. The University of Chicago Press.

Latour, B., 1987. Science in Action. Harvard University Press, Cambridge, Massachusetts, London.

Leonelli, S., 2014. What difference does quantity make? On the epistemology of Big Data in biology. Big Data & Society 1 (1), 1–11.

Marcum, J.A., 2017. Clinical decision-making, gender bias. Virtue Epistemology, and Quality Healthcare 36 (3), 501–508.

Madsen, A.K., Munk, A.K., 2019. Experiments with a data-public: moving digital methods into critical proximity with political practice. Big Data & Society 6 (1), 1–14.

Mayer-Schönberger, V., Cukier, K., 2014. Big Data: A Revolution that Will Transform How We Live Work, and Think. Mariner Books, Houghton Mifflin Harcourt.

Moats, D., Seaver, N., 2019. "You social scientists love mind games": experimenting in the "divide" between data science and critical algorithm studies. Big Data & Society 6 (1), 1–11.

Mortelmans, D., 2007. Handboek Kwalitatieve Onderzoeksmethoden. Acco.

Murdoch, T.B., Detsky, A.S., 2013. The inevitable application of big data to healthcare. J. Am. Med. Assoc.: JAMA, J. Am. Med. Assoc. 309 (13), 1351–1352.

Oldenhof, L.E., 2015. The Multiple Middle: Managing in Healthcare. Erasmus University, Rotterdam.

PacMed, 2019. Projects. Retrieved from [September 2019]. https://pacmed.ai/nl/projects/.

Paul, H., 2011. Performing history: how historical scholarship is shaped by epistemic virtues. Hist. Theor. 50 (1), 1–19.

Pigliucci, M., 2017. The virtuous skeptic. Skeptical inquirer. The Magazine For Science and Reason 41 (2).

Rathenau Institute, 2018. Science in Figures: Infrastructure of Knowledge. Retrieved from [March 2020]. https://www.rathenau.nl/en/science-figures/policy-and-structure/infrastructure-knowledge.

Ribes, D., 2019. STS, meet data science, once again. Sci. Technol. Hum. Val. 44 (3), 514–539.

Rose, N.S., 2007. Politics of Life Itself: Biomedicine, Power and Subjectivity in the Twenty-First Century. Princeton University Press.

Rüppel, J., Voigt, T.H., 2019. The death of the clinic? Emerging biotechnologies and the reconfiguration of mental health. Sci. Technol. Hum. Val. 44 (4), 567–580.

Sample, I., 2018. 'It's Going to Create a Revolution': How AI Is Transforming the NHS. The Guardian.

Smits, M., 2002. Monsterbezwering. Boom uitgevers, Amsterdam.

Somsashekhar, S.P., Sepúlveda, M.J., Norden, A.D., Rauthan, A., Arun, K., Patil, P., Ethadka, R.Y., Kumar, R.C., 2017. Early experience with IBM Watson for Oncology (WFO) cognitive computing system for lung and colorectal cancer treatment. J. Clin. Oncol. 35 (15_Suppl. l) 8527-8527.

Star, S.L., Griesemer, J.R., 1989. Institutional ecology, 'translations' and boundary objects: amateurs and professionals in berkeley's museum of vertebrate zoology, 1907-39. Soc. Stud. Sci. 19 (3), 387–420.

Stevens, M.J., Wehrens, R., de Bont, A., 2018. Conceptualizations of Big Data and their epistemological claims in healthcare: a discourse analysis. Big Data & Society 1–21 0 (0).

Timmermans, S., Berg, M., 2003. The Gold Standard: the Challenge of Evidence-Based Medicine and Standardization in Health Care. Temple University Press, PA.

Timmermans, S., Tavory, I., 2012. Theory construction in qualitative research: from grounded theory to abductive analysis. Socio. Theor. 30 (3), 167–186.

Vaughan, D., 1999. The role of the organization in the production of techno-scientific knowledge. Soc. Stud. Sci. 29 (6), 913–943.

Van Lente, H., 2012. Navigating foresight in a sea of expectations: lessons from the sociology of expectations. Technol. Anal. Strat. Manag. 24 (8), 769–782.

Vidal, F., Ortega, F., 2017. Being brains: making the cerebral subject. In: Forms of Living. Fordham University Press.

Wehrens, R., 2015. The potential of the imitation game method in exploring healthcare professionals' understanding of the lived experiences and practical challenges of chronically ill patients. Health Care Anal. 23, 253–271.

Ziewitz, M., 2017. A not quite random walk: experimenting with the ethnomethods of the algorithm. Big Data & Society 4 (2), 1–13.