Clinical Trials and the Drive to Material Standardisation: 'Extending the Rails' or Reinventing the Wheel?

Catherine M. Montgomery

University of Oxford, UK / catherine.montgomery@phc.ox.ac.uk

Abstract

There have long been calls from within both industry and academic groups to reduce the bureaucratisation of clinical trials and make them more 'sensible', with the focus on approvals and guidelines. Here, I focus on the mundane environments of a multi-centre clinical trial to ask how 'sensible' it is to standardise trials at the level of material objects. Drawing on ethnographic data collected in the UK, South Africa and Vietnam, I present three vignettes of material standardisation. While acknowledging some positive effects, I argue that standardising in this way may be antithetical to sustainable and relevant clinical research. Three dimensions of this are discussed: 1) the external validity of evidence from pragmatic trials 2) the gap between experimentation and implementation and 3) long-term site capacity to conduct research. Drawing on the literature on 'situated standardisation', the paper concludes by suggesting a greater acknowledgement of the need for trials not only to be 'sensible' but also 'situated'.

Keywords: clinical trials, standardisation, materiality

Introduction

The lure of standardisation exerts a firm grip on clinical research. Over the twentieth century, increasing moves have been made to modernise, standardise and thereby cement medicine's status as 'science'. These have included the organisation and regulation of pharmaceutical drug trials and the rapid ascent of the evidence-based medicine movement (see Epstein, 2007). As the poster child of evidence based medicine, the randomised controlled trial (RCT) relies on standardisation to make experiments consistent, transparent and comparable. Various kinds of standard are at work, and here Timmermans and Epstein's (2010) sub-types are useful: terminological standards work to stabilise meaning across different sites and times and enable individual elements to be aggregated into larger wholes; procedural standards delineate how processes are to be performed; and design standards 'set structural specifications: they define the properties and features of tools and products' (Timmermans and Epstein, 2010: 72). Such standards ensure the uniformity and mutual compatibility of sociotechnical systems (see also Timmermanns and Berg, 2003) and have become a defining feature of the way in which evidence is produced in clinical trials.

In multicentre clinical trials, as in other domains of social life, standards mean that people and things can be coordinated in a consistent and measurable way that would otherwise be almost impossible to achieve. Standard Operating Procedures (SOPs), Working Practice Documents, technical manuals and protocols, to name but a few of the everyday tools of clinical research, allow communication between different (potentially incompatible) systems, and enable movement, calculation, precision, universality and objectivity across sites. However, standards, being the result of social work (Bowker and Star, 1999; Lampland and Star, 2009) and requiring collective labour, negotiation and buy-in from multiple stakeholders (Berg, 1997; Fujimura, 1992; Jordan and Lynch, 1998), are not politically neutral. As Timmermans and Epstein (2010: 74) aptly put it, their "objectivity, universality, and optimality are hard won victories that can be heavily contested by third parties lobbing accusations of bias and politicization".

In relation to clinical trials, then, it is important to ask whose benefits are served by particular standards, and in the case of conflicting standards, whose should prevail. While the standardisation of clinical guidelines has been well analysed in this respect (Cambrosio et al., 2006; Castel, 2009; Knaapen, 2013; Knaapen et al., 2010), a feature of multi-centre clinical trials that has been relatively neglected is the standardisation of material settings. Perhaps because of their mundanity these have received little or no attention in the clinical trials literature and limited analysis in social studies of science.

A notable exception is Petty and Heimer's (2011) analysis of how HIV clinical research shapes clinics not only at the point that research findings are implemented, but through the very process of conducting such research in the first place. Petty and Heimer argue that clinical trials transform medical practice in the places where they are conducted by modifying the material environment, reorganising bureaucratic relations and increasing the valorisation of research. Clinics that *conduct* research, they argue, are in a better position to *implement* the results of research because they have been re-made in ways that smooth the transition of practice. They draw on Latour's (1983) image of scientific facts being like trains that that do not work off their rails to characterise this phenomenon as "'extend[ing] the rails' that allow scientific research results to be driven into the clinic" (Petty and Heimer, 2011: 357).

Like Petty and Heimer (2011), I am interested in global public health trials, which tend to fall at the 'pragmatic' end of the explanatory-pragmatic continuum (Thorpe et al., 2009). That is, they seek to determine the effects of an intervention under *usual* rather than *ideal* conditions. With the emphasis on usual conditions, the assumption might be that standardisation of the experimental environment across sites is minimal or even absent. However, trial sponsors often invest in site infrastructure and provide standardised consumables (such as diagnostic tests, laboratory reagents, and stationary) as part of scientific 'capacity building'. It is this material standardisation with which the current paper is concerned. In contrast to Petty and Heimer's (2011) conclusion that the standardisation of research environments leads to the more ready adoption of new research findings in medical practice, I start from the more sceptical position that the proliferation of standards in research sites may have negative as well as positive effects. This position is informed by ten years working on medical research projects and frequently hearing sites referred to as vessels through which clinical trial traffic can be routed. As trials from a range of sectors, sponsors and disease areas proliferate at 'good' sites, how 'sensible' does trial-specific standardisation become, and what are its effects?

For a number of years there have been calls for an end to the bureaucratisation of clinical trials (Groves, 2009; Shurlock, 2013). This is evident in the re-writing of the European Commission's guidelines on the conduct of clinical trials, the Sensible Guidelines Group (SGG) and the US Clinical Trials Transformation Initiative. At the heart of these initiatives is a belief that over-regulation has brought the development and testing of new pharmaceuticals to its knees and is delaying the introduction of potentially life-saving drugs to patients around the world. A 2008 special issue of *Clinical Trials* brought together papers written by members of the SGG which focused on a range of aspects of such over-regu-

lation, including the restrictive interpretation of privacy laws and its negative impact on the use of personal health information in trials (Armitage et al., 2008); excessive and inefficient on-site monitoring (Baigent et al., 2008); obstacles to conducting trials with vulnerable patient populations (Cook et al., 2008); and overuse of and overly complex adjudication of clinical events (Granger et al., 2008).

An overarching critique made by the movement towards sensible clinical trials is the inefficiency of a 'one-size-fits-all' approach to conducting and regulating medical research. Proponents question how appropriate, relevant and representative the regulatory guidelines for the conduct of RCTs are (Yusuf et al., 2008). The International Committee of Harmonisation's guidelines to Good Clinical Practice (ICH-GCP) have come under particular fire, in particular their uncritical application to all kinds of clinical research study in all kinds of differently-resourced settings (Lang et al., 2011). Rather than taking the guidelines as a set of principles by which to conduct ethical trials and report accurate and reliable data, study sponsors have instead "appl[ied] the guidelines as a single standard" (Lang et al., 2011: 1555).

The movement towards sensible clinical trials provides an interesting starting point for considering standardisation in clinical research, by virtue of the fact that it provides a ready-made critique from within the orbit of the pharmaceutical industry itself. However, the SGG, it should be noted, sits in a field dominated by a powerful pro-pharmaceutical industry lobby, which has sought to de-regulate trials as part of more systematic neoliberal attempts to make pharmaceutical regulation less restrictive. As Abraham, Davis and others argue, this has produced toxic results for public health (Martin et al., 2006; Davis and Abraham, 2013; Light and Lexchin, 2012). Moving the focus to trial practices themselves, the literature on the political economy of clinical trials could not be clearer about the links between pharmaceutical neoliberalisation and the enrolment of vulnerable populations into potentially exploitative regimes of commercial experimentation (Fisher, 2009; Petryna, 2009; Sunder Rajan, 2007).

In this paper, I examine the material standardisation of trials and – as a way to contribute both to debates in STS and in Global Health – ask the question, "how sensible are clinical trials?" In adopting this terminology, I am clearly not condoning the deregulation of clinical research, or celebrating a more permissive approach to pharmaceutical experimentation; instead, my aim is to challenge what is done in the name of good science, where this indexes both the rigour of the experiment and its moral claims to build capacity. I use the term 'sensible' heuristically to frame the analysis of a publicly-funded academic-run trial into anti-Tuberculosis drugs, an area in which so-called market failure has led to the need for new regimens to be tested outside the industry model (Cousins, 2016; Frick, 2016). Within this context of sparse investment and infrastructural poverty, the question as to how 'sensible' each step in the clinical testing process was was never far away. By bringing a classic STS concern (standardisation) to bear on the topic of global health trials, I aim to craft an analysis that – in line with the aims of this special issue – speaks critically but constructively to both fields.

The paper proceeds as follows. Firstly, the ethnographic study of a specific multi-centre trial is presented. Thereafter, I go on to present three vignettes about material standardisation: diagnostic devices, drugs and samples. The vignettes are analysed to show the social effects of standardisation for the sites taking part in the trial. In the ensuing discussion, I argue that while material standardisation can have positive effects, it may also contribute to conditions which are antithetical to sustainable and relevant clinical research. I discuss three dimensions of this, namely 1) the external validity of evidence from pragmatic trials 2) the gap between experimentation and implementation on which a new policy's success can depend and 3) long-term site capacity to conduct research. Drawing on the literature on 'situated standardisation' (Zuiderent-Jerak, 2007; Engel and Zeiss, 2014), I conclude by suggesting a greater acknowledgement of the need for trials not only to be 'sensible' but also 'situated'.

Methods

From May 2012, for a period of 12 months, an ethnographic study was undertaken to understand

the processes of governance, exchange, sharing, value-creation and appropriation in a transnational biomedical research project. The study was designed to investigate how different partners in a multi-sited trial conceptualise and attribute meaning to collaboration; how the division of labour is organised; how exchange practices, such as sharing, giving, and transferring shape the trial network; and how transactions occur, e.g. in relation to collaboration, training, recruiting participants, sharing information and materials. In order to do this, I conducted twelve months' observation of a publicly-funded RCT investigating new treatment options for multi-drug resistant tuberculosis (MDR-TB), basing myself at the trial's UK coordinating site, and attending over 30 team meetings and teleconferences. I accompanied UK staff on site initiation visits to two clinical sites in South Africa and Vietnam (each lasting two to four days), as well as visiting these sites independently to observe the scientific and administrative practices of the trial (5 weeks in total). I conducted a total of 34 interviews with staff working on the trial across coordinating and clinical sites, including principal investigators, trial sponsor, trial managers, statisticians, clinicians, nurses, pharmacists, and monitors.

During the observation, fieldnotes were taken and these were typed up on a daily basis. Interviews were digitally recorded, transcribed, and translated where necessary. All data were imported into NVivo software for qualitative data analysis. Coding, memo-writing and interpretation followed the principles of constructivist grounded theory (Charmaz, 2006). An initial process of detailed line by line coding within interviews and fieldnotes led to the development of a set of provisional categories, used to code subsequent transcripts in a more focused manner. This iterative process involved testing the adequacy of categories against the data (constantly turning between codes and data) and then of moving between cases (comparing data to data).

Ethics approval was obtained in the UK by the University of Oxford Central University Research Ethics Committee; in South Africa and Vietnam by the relevant local institutional review boards; and in addition by the International Union Against Tuberculosis and Lung Disease Ethics Advisory Group. For institutional observation, written informed consent was obtained at each site from the principal investigator, and verbal consent from staff. All staff gave written informed consent prior to being interviewed.

Study settings

The three sites which formed part of this multi-site ethnography consisted of a publicly funded UK clinical trials unit, a government-run TB referral hospital in South Africa and a government-run tertiary hospital in Vietnam. In South Africa, the TB hospital encompassed, within the hospital grounds, a research unit separately funded and staffed by research grants, but recruiting patients from the general hospital population. A clear demarcation between research and care was noted by research staff, with some antagonism reportedly created by the different sources of funding, which manifested itself in different nursing duties, work uniforms, prestige, and so on. By contrast, in Vietnam, research was conducted by government-funded staff and was seen to be an integral part of the career trajectory for those who wanted to progress. Research was valued by hospital management and, while creating more work for staff, was also seen as a conduit to changes in patient management based on the latest evidence. The sites in question were just three of a larger number of sites taking part in this multicentre trial across Europe, Africa and Asia. They were selected for the ethnographic study in consultation with the trial management group, trial steering committee and the principal investigators of the sites themselves. The trial and its sites are not named in this paper in order to preserve anonymity.

Findings

Standardising 'usual conditions'

> There's a quick examination of little sealable clear plastic bags that the study drugs will be dispensed in; also of how to print off labels for said bags. The pharmacist shows us how she prints off labels and cuts them to size with a ruler and a craft knife.
> The visitors say that in this trial, she will be able to dispense with this time-consuming task, because they will provide a special printer and labels that are the right size. The pharmacist looks nonplussed;

> I can see she already has two printers in her small office, a HP and a Cannon. Adding a third printer for the study, the Zebra printer, seems almost comical. – Extract from fieldnotes, site visit to the hospital pharmacy, Vietnam

Within the scientific community, the trial was regarded foremost as a pragmatic, rather than an explanatory trial. The trial in question asked, "does this intervention work under usual conditions?" and did so in a number of different settings in order for the results to be seen to be widely (globally) applicable and therefore to form the basis for a World Health Organisation (WHO) recommendation for MDR-TB treatment. According to the protocol, the trial explicitly adopted a practical, programme-based design to make sure that if successful, the results would be generalisable to routine programme settings.

The distinction between explanatory and pragmatic is important here because it indicates an intention on the part of trialists to either create and control an 'ideal' environment for the experiment or to refrain from intervening in the experimental setting and let events take their course. However, what became clear during the ethnography – and is illustrated by the fieldnote extract above – is that even at the pragmatic end of the spectrum, considerable effort goes into procuring materials for trials, from pharmaceuticals to medical equipment to stationary. In the trial in question, a multitude of objects were shipped to the clinics in order to ensure material standardisation. Such objects included ziploc bags, drugs in the intervention arm, electrocardiograph (ECG) machines, label printers, printed CRFs and logs for recording everything from the temperature in the drug store to sample chain of custody.

The trial also standardised the way in which data was collected, not only in terms of the physical forms just mentioned, but also in the structure of questionnaires, the phrasing of questions, the units of measurement and so on. The questions assumed a single reality in multiple settings that could be apprehended by asking the same question the same way in different places and at different times. A variety of texts instructed those implementing the trial in how it was to be achieved, and, additionally, regulated this. A prime example was the protocol; others included SOPs and Working Practice Documents. These texts were strictly controlled: their circulation was limited and any changes had to be made through a centralised and audited process. 'Version control' was observed, to ensure old versions of the text were not in use. In some cases, an electronic infrastructure was in place both to govern and to provide an audit trail of changes to texts. This is part of a much larger regulatory framework mandated by the Medicines and Healthcare products Regulatory Agency (MHRA) which oversees the activities of clinical trials units in the UK.

Below, I present three examples of material standardisation at work.

Diagnostic devices

Electrocardiography literally involves the inscription of the electrical activity of the heart. It is used as a diagnostic device to measure abnormal heart rhythms, be this during regular patient care or specifically for research purposes. As part of safety monitoring, ECGs were used in the trial to monitor patients' heart activity and reduce the risk of adverse events related to one of the study drugs. Additionally, ECG data were being collected to assess the impact of study drugs on QT (the QT interval is the time from the start of the Q wave to the end of the T wave and represents the time taken for ventricular depolarisation and repolarisation), for which there was little existing data. ECG machines were purchased centrally and shipped out to trial clinics. The shipping process entailed many delays and significant labour from the coordinating team to ensure safe and timely delivery. Indeed, much time was devoted to discussing the ECG machines in team meetings and teleconferences. The clinics could not start enrolling patients into the trial until they had received the ECG machines, been trained in their use, and got them set-up and working. This put all involved under a certain amount of pressure, since budgets were being spent employing staff for the trial, and these budgets had to last until the trial was completed.

ECG was one means through which the trial participants' bodies were translated into data; the beating heart 'travelled' from the trial clinics to the UK and back again through a circuit, coordi-

nated not by a clinician, but by a data manager. During a presentation at one trial site, the UK data manager elaborated on how this circuit worked: the ECG machine – a MAC800 – is shipped from London, via the trial sponsor in Paris, to Johannesburg. A doctor or nurse at the site uses the ECG machine to take readings from the patient's heart. The MAC800 comes with software which enables the medical staff who conduct the procedure to save the output as a PDF and email it to the data manager at the coordinating site in the UK. The coordinating site then passes the data on to an independent cardiologist at a UK university, who has offered his expertise to the trial. He interprets the readings and can make recommendations on changing the patient's treatment, if necessary. So it is that a beating heart in South Africa is transformed and travels all the way to the UK.

Why was it, I asked a clinical investigator in the UK, that the ECG machines needed to be shipped around the world; weren't such things available locally?

> The ECG machine, now that was a standardisation issue and that was, well …a) they wouldn't have them anyway and we wanted the reports to be the same on each, and set them up so that they could get the same information out and that was standardised. Plus that it was set up for them and they didn't have to each figure out how to get out the necessary information on whatever system they happened to have bought.

She went on:

> R: [The independent cardiologist] advised on our approach to the monitoring of these patients and how we approach that in the protocol. So he advised us on that, not just individual cases. He advised us on what sort of ECG machines and what sort of holters we needed.
> I: The MAC800?
> R: He didn't ask for it that specifically, but you know, we said how difficult it would be, because we're going to all these places that aren't used to doing, haven't done ECGs for years and he said, "Oh you just get one of these machines that, you know, that print out the answer for you".

The ECG was thus very much a part of the evidence-making apparatus of the trial. It was of central importance that the readings be standardised across different settings, and it was seen as desirable that the machines simply 'give the answer' rather than requiring extensive staff training. The schedule for conducting ECGs was also standardised in the protocol; to avoid confusion in the data, sites were asked not to conduct ECGs unless scheduled.

While the ECG machine has potential as a knowledge tool in clinical practice generally, this potential was foreclosed by the way it was configured in the research. In the hospital in Vietnam, I was told that the ECG machine was the only one available on the ward, but that it could only be used for trial participants because the software was programmed to require a study number. If a patient who was not enrolled in the trial needed an ECG, the staff had to borrow a machine from the emergency resuscitation department:

> For the ECG machine, it's required that we have patient information; it's like a key. If the machine doesn't get that information, there's no key and it can't measure. Some doctors ask me to measure their patients and I say that if there's no trial code, the machine won't work. Therefore it can't measure, so we've never used it to measure non-trial patients.

Therefore, whilst the UK investigators aimed to help the site clinicians by simplifying the technology, in practice, this meant that site staff were effectively locked-out of the machine, unable to adapt its use to their local requirements. The ECG machine delimited the experimental context; it only worked on and for certain people who had codes (study number, user code) and whose relationship with the machine and with each other was directed to the experimental goal. This was likewise reflected by staff in South Africa, as the following extract illustrates:

> I: Do you still do the ECGs [in spite of the fact that government staff provide routine care]?
> R: Yes. Because ours is a different ECG. It has to be saved onto a disk and it must be emailed to our [trial] data team. So that's also a procedure on its own.

While staff at both sites were excited to receive new, high-tech equipment, because the machine was standardised across trial sites, it was difficult for them to incorporate its use into their routine care setting. In some cases, this disrupted established relationships between local staff working on the trial and their colleagues.

Drugs

> When you deal with a clinical trial, the minimum you can do for the benefit of patients entering the trial – you ensure the quality of the medicines, but also there is even a research objective, where you need to ensure that there will be a repeatability of the results of what you are assessing in terms of regimen for the drugs provided to South African sites, for Vietnamese sites…If you've got different qualities, what kind of assessment will you make in the end? It will be completely unhomogenous. Which is not at all what you look for in a clinical study. – Coordinating site staff member

The trial was testing a new regimen of existing drugs which had been approved, licensed for various indications, and were already on the market. The novelty lay in putting them together in a particular combination of dosages and schedules. Patients randomised to the intervention arm received this new regimen, while patients randomised to the control arm received the locally-used WHO-approved MDR-TB regimen. Neither patients nor clinicians were blinded to treatment allocation, but laboratory staff, who produced results on patient outcome measures, were.

The drugs for the trial regimen were standardised. They were purchased through the Global Drug Facility for TB, a WHO-housed procurement mechanism that ensures a single quality standard. The drugs for the *control* regimen were not standardised between the different trial sites, since each country has its own National TB Control Programme with its own protocols and procurement channels. The procurement of standardised drugs across the trial was one of the biggest challenges the trialists faced, since it involved estimating timelines for drug dispensation and expiry, obtaining import permits and VAT exemptions, negotiating delays in customs, acquiring and monitoring suitable storage facilities, etc. In fact, since the drugs the trial was testing were all already licensed, they were available in-country, but because the quality could not be assured, the decision was taken to import them.

As the staff member quoted above notes, standardising the drugs in the trial does two things: it ensures the patient receives a quality product and it allows a comparison to be made across different settings as to the effectiveness of the new regimen. The trialists seek pharmaceutical homogeneity in order to conduct a rigorous scientific experiment; in effect, their aim is to ensure that the drugs in the different countries are all the same.

The drugs procured for the intervention arm in the trial were not treated the same way as the drugs in the control arm (the WHO approved in-country regimen). The imported drugs had to be stored separately from other drugs, and the hospital pharmacies had to create special spaces for this. A raft of paperwork was associated with dispensation, swallowing, return and destruction. Logs had to be completed for accounting purposes (prescription register, receipt log, packing log, return log, destruction log…); drugs had to be dispensed into individual daily and weekly plastic bags (to ensure consistency and correct dosing); and any un-swallowed drugs had to be returned to the pharmacy (to enable audit and prevent circulation on the black market). It was not just the drugs themselves that became differentiated in this way; the people handling them also acted and were acted upon differently: patients were marked out as different by receiving their pills in individual bags rather than straight off the dispensing trolley; nurses had to handle returns differently, keeping all un-swallowed tablets in their bags and sequestering them for accounting purposes; pharmacists and coordinators had to destroy drugs, which normally would be reintroduced into circulation.

The destruction of drugs was a contentious issue, which was not well understood by all staff, as the following conversation with the pharmacist at one site illustrates:

> I: In terms of the destruction of drugs, can you tell me how that works?
> R: For us, for this trial, it's mostly been patient returns, and obviously once a drug is expired, it will also go onto the destruction. So when the patient

> brings back returns it's written into the destruction log and placed into a green bin, as dedicated for destruction, so when that reaches a certain level, we'll get authorisation from [the UK trial manager] and it will be uplifted and an outside company takes it off for destruction and then they'll give us a destruction certificate.
> I: And why do you have to seek permission to get the drugs destroyed?
> R: I don't know [laughs]. It's written in the log there! "Permission granted" – I don't know why!

In a more reflective moment, a coordinator of the National TB Control Programme in one country told me, "I know it's research, but I still cannot understand why we have to destroy drugs… The problem here is that the drugs for today are exactly the same as the drugs for tomorrow…why don't we reuse them? If the dosages are different then I'm fine with not using the drugs again, but they are all the same, so why do we throw them away?" While the drugs looked the same to this doctor, who was concerned with treating patients as an end in itself, to the trialists – who were concerned with treating patients as a means to answer a scientific question – they were not. Today's dose may have been the same as tomorrow's dose chemically, but it was not in evidential terms. In order to capture knowledge of how well the drugs are working, it is important to the trialists to know how many of the dispensed drugs have been taken. Since it seems feasible that the trial implementers could simply write down the number of un-swallowed drugs before re-dispensing them, one is led to ask what the sequestering and destruction of drugs in the trial achieved.

The procurement pharmacist told me there were two reasons. The first was to make sure that none of the un-swallowed drugs made their way onto the black market, where their use could not be controlled:

> [In] all these countries, withdrawal of expired medicine is very poorly controlled…so it's very tempting to do some black market just with whatever, even the expired drugs, because a lot of people don't know that expired drugs could do harm or could not be efficient, so they will find customers for it …So just to make sure that bringing extra drugs in these countries, at least we are responsible for how they're going to be destroyed, just to make sure they're not going to nourish any dirty system.

The second was to satisfy regulatory audit, in which the sponsor could be asked to back up the trial result with evidence from pharmacy accounting logs:

> I used to be a Good Clinical Practice auditor and… there is a very easy way to find out whether a company or an entity – a sponsor which has organised a clinical trial and is announcing outcomes on 350 patients – to make sure that really 350 patients have been treated. There is a very easy way to do it: "OK, give me all your files with how many drugs were provided, how many were dispensed, and which were destroyed." And if the balance doesn't match…Mmhmh! And often people, when they want to cheat, actually it's very difficult to really set up false drug dispensary forms.

The first reason the pharmacist gives concerns the physical status of the drug as chemical; the second, its status as evidence. In the former, the sponsor is imagined as a responsible actor in the local economy of pharmaceutical dispensing: trial drugs must not enter the informal marketplace and must therefore be destroyed. In the latter, it is portrayed as a responsible actor in the global economy of evidence-making: the destruction of trial drugs makes accounting practices add up. The practice of destroying drugs rests on a belief that records from actual drug destruction can be differentiated from faked records. It is very difficult to cheat, the pharmacist says. This belief privileges empiricism by implying that data derived from direct observation can be identified as true over made-up data. The obvious truth of the data lies in its correspondence to real events. Therefore, in order to achieve convincing evidentiary ends, the drugs must actually be destroyed.

While this position may be sensible from a drug regulatory perspective, from the perspective of some of the people working in the clinics, it was incomprehensible. Local staff understood that they must operate according to global standards, written into guidelines they had been given, but did not necessarily understand the reasons for the guidelines. In the resource-limited settings of this trial, where there were insufficient drugs to treat

patients, the idea of 'throwing away' good drugs was seen by some to be uneconomical, wasteful and morally wrong. This was compounded by the fact that achieving drug destruction according to global standards significantly increased the burden of work for public sector staff, who already had a heavy workload.

Laboratory results

As part of the national TB management protocol in South Africa, patients must provide sputum samples to be sent to the National Health Laboratory Service (NHLS). Samples take several weeks to be cultured and results are returned to the ward to inform patient management. At the same time, however, patients enrolled in the trial had to provide a second sample, which was sent to a specialist research laboratory, where the same information was extracted, but based on a set of laboratory procedures that were standardised for all the trial sites to follow. This information was also returned to the ward, usually more quickly than the NHLS results, and was stored in the patient's file. Following GCP, results did not contain patients' names but only study numbers.

A staff member described the tension that this standardised procedure produced in the hospital:

> We have to send our bloods through an accredited laboratory and our sputum specimens through an accredited laboratory. *So we have to have a parallel process*…To be a research-accredited lab, you need certain standards, I don't even know what they are…Sponsors want standardisation across all the sites, so therefore they select specific laboratories to do the work. And they accredit them. Now one of the ways that this has posed a challenge is that our bloods and our sputum specimens don't have patients' names on them; they've got study numbers and ID numbers. And for example we had a patient with a low potassium and we spoke to the doctor, but she was concerned that this piece of paper didn't refer to this patient because it didn't have the patient's name on. (emphasis added)

In spite of the pragmatic nature of the trial, and the attempt to produce evidence under usual programme conditions, it is clear that in relation to various parts of the care cycle, 'parallel processes' are instituted in order to standardise. In relation to laboratory tests, the knock-on effect was felt acutely at the patient-provider interface, as the following quote illustrates:

> Well I know from the nurses that there's sometimes some antagonism because obviously two doctors trying to manage a patient can cause problems, and especially since our results come from a different lab and their results come from NHLS labs … so the ward staff are then a little bit reluctant to react on results where it doesn't actually have the patient's name on it. They want their own results to come through.

As this example illustrates, when research is introduced into a routine care setting, different sets of standards may clash. Standardisation for care and research do not necessarily map onto each other; new forms of standardisation (such as the processing of lab results) can be interpreted as a de-valuing of existing practice and a critique of current standards. This can cause resentment among care staff, with a potential knock-on effect for patients and subsequently for recruitment and retention in the trial. Various examples of this were given by trial staff, perhaps one of the most notable being the delay to patients' discharge from hospital following a negative smear result:

> We've had a couple of occasions when our smears come back as negative on a Thursday and the patient wants to go home for the weekend, but the hospital smear is not back. So now the patients say, "But study smear's negative, why can't I go home?" Or we get a culture result that comes back negative, theirs is still pending, and it can be a two week difference, which …for our patients it's very significant.

Extending the rails or reinventing the wheel?

Because things aren't standard across different settings, trials tend to standardise. But because things aren't standard across different trials, standards proliferate. The coordinating staff frequently had to negotiate the practicalities of standardising across multiple settings, being well aware of the scientific and regulatory requirements placed on clinical trials, but equally cognizant of the very real effects for clinical sites of any attempt to standardise practice. This was vocalised as a desire not to

'reinvent the wheel' where local guidelines and practices could be accommodated within the protocol, but equally as a responsibility to maintain control over unchecked heterogeneity.

For clinical sites, exposure to standardisation through multicentre trial participation can be positive, in that a range of new skills and commodities can be acquired by working with different collaborators on different experiments. At both clinical sites, staff spoke highly of the benefits of participating in international research projects, including the trial in question. Reported benefits spanned the levels of the institution, patients and staff. At the institutional level, multicentre trials could lead to reputational gains, by virtue of being associated with well-known international partners. At the patient level, benefits related to early access to new treatment standards. And at the staff level, interviewees spoke of acquiring greater knowledge and skills and being able to work to international standards. For example:

> If we sign African investigators on and train them how to do research in [this trial] and then they go on and write investigator-initiated studies because of the experience they had in [this trial], that's capacity building. You know GCP training and all that kind of stuff.

> I think that the international research is helpful not only for patients but for staff also. We learn so much when we run this study. Because for international research, we have to have many standards and these are international standards and in Vietnam we have not so much experience. But step by step, by studies, we learn and hope we can reach the international standards.

However, working on multicentre trials was also said to have negative effects, such as staff being overwhelmed by the continual need to learn and implement new processes; duplication of systems and equipment that was not needed (see the printer example at the start); and fractures in the transferability of research findings to the standard of care once a given research study yields results. Criticising the duplication of work needed between the databases of the National TB Programme and the trial, one staff member commented:

> This study goes its own way, different forms and documents and templates, it does create a burden for staff. If [this trial] is successful and later on it still requires procedures like that, I don't think they can follow it.

Another at a different site alluded to frustrations with parallel processes and the lack of agency site staff had in relation to their everyday working practices:

> Well we did try at the beginning of the study to ask [the UK trial manager] to give us a log so that we could amend it so that it would be in line with our working practices and we don't change too much of our SOPs and things like that. Because it's new also, it would be difficult to remember what you have to do. So they were reluctant to send us the format in which we could change it. They said they would only send it to us in pdf format and then they'd ask us "what is the change that you need and give us an explanation as to why we need to change it".

These examples show the tensions between standardisation and localisation that must continually be negotiated by all parties in an international multicentre trial. The material artefacts mentioned in this paper (diagnostic medical equipment, databases, drugs, printers, stationary, sputum specimens, and test results) necessitate a raft of novel practices, and it is often these practices – rather than the artefacts *per se* – which result in the intangible gains touted for transnational research, i.e. capacity building, reputation, etc. As currently conceived, the two cannot be separated. The challenge for global health trials is to acknowledge the value in local specificity and rather than seeking to efface it, work with it to produce science that is both rigorous and situated.

Discussion

How 'sensible' is standardisation? In this paper, I have provided three examples of material standardisation in a multicentre clinical trial which trouble an easy answer to this question. On the one hand, clear rationales exist for such standardisation on ethical, regulatory and scientific grounds. On the other hand, efforts to make things the same across diverse care environments, even

down to the level of how drug labels are printed, potentially create conditions which are antithetical to sustainable and relevant clinical research, including 1) the external validity of evidence from pragmatic trials 2) the gap between experimentation and implementation on which a new policy's success can depend and 3) long-term site capacity to conduct research. I address each of these in turn below.

The first dimension of the sensibleness of standardisation concerns the extent to which the experimental environment reflects usual conditions. As each new trial imports (self-) standardised diagnostic equipment, drugs, stationary, and software, clinical sites become repositories of both material worth and evidentiary values. When a trial ends, the result is extracted, while these ordinary yet un-usual objects remain behind. As each trial requires and produces its own uniformity, and as the traffic in trials increases (particularly in 'good' sites), so the 'representative care setting' is re-invented again and again. This leads to questions about what the site is ultimately representative of, and how far the evidence produced within it can easily be adopted or transferred to other places. In contrast to Petty and Heimer's work suggesting that a clinic's ability to conduct research positively influences its ability to implement research findings by 'extending the rails', I have argued that the lack of a stable organisational environment can have adverse effects in this respect, leading to the continual reinvention of the wheel. While Petty and Heimer (2011) acknowledge that the extent and permanency of standardised practices in clinical research settings can vary, they nonetheless emphasise the positive association they engender between research and implementation. I agree with them that "conducting research is likely to have its most lasting effects when the network of ties and the infrastructure built and reconfigured in the course of doing a research project are later appropriated by subsequent research projects and care programs" (Petty and Heimer, 2011: 357), but based on over a decade of observation, this is not commonly the case.

Entangled with this first dimension is the second, which concerns the bifurcating effect that material standardisation can have on the relationship between research and care. As the empirical examples in this paper illustrate, standards change the worlds upon which they are imposed, reconfiguring relationships between colleagues and between medical staff and patients. While such changes can be positive, arguably driving up professionalisation and introducing the latest standards into clinics, they can also result in a 'double standard', where existing local practices co-exist alongside the new and 'universal', but are devalued. The privileging of external solutions over what is available locally can have unanticipated effects, as numerous studies of technology transfer have shown (see e.g. Müller-Rockstroh, 2012).

What's more, whereas extensive labour is required by trial managers to circulate standardised materials during the course of an experiment, this labour (and the artefacts involved) is usually omitted from scientific accounts of the results. It is troubling that where the aim of a trial is to usher in a new 'standard of care' (as was the case with the trial described in this paper), the standard of care is imagined in pharmaceutical terms rather than as the sum of social and material relations which have brought the result about. Marks' (1997) observation that "even the simplest RCT is the product of a negotiated social order, replete with decisions…and with unexamined assumptions" could only be more pertinent had he added 'mundane artefacts' to the list. While pragmatic trials are designed to generate 'real world evidence' for clinical decision-making in a valid way (Zuidgeest et al., 2017), what the real world is made of is a question needing more granular and transparent treatment.

The third dimension concerns the temporality of standardisation and related to this, the sustainability of 'capacity building' when this is used as a proxy. What is the relationship between standardisation and sustainability in clinical research? It would be logical to assume that once a site is up and running with its GCP and its SOPs, it would be set up to run all future trials, or at least future trials of a similar nature. However, this is not necessarily the case, because each new trial that comes along still purchases and ships in new equipment, new forms, and a variety of other 'new' standards embodied in everyday objects. But what would

happen if these objects were not shipped out and the experiment was done using what was already available locally? If we applied this thought experiment to the real experiment discussed in this paper, the pharmacist would continue to use the printer she has and cut out drug labels according to her own method. The coordinating team would accept that the pharmacy already has systems in place and not ship reams of new logs to be completed in addition to those that are being completed already. ECG machines would be purchased locally and made available for staff to use according to local needs. The results would not be standard, but the information would be there nonetheless.

When things aren't standardised, the adjustment, the interpretation, the making things fit together, has to happen on the part of a trial's coordinating team as well as on the part of the sites. What is privileged is not uniformity but variability, availability, suitability, sustainability. Such an approach would require a greater recognition of the situatedness of trial results, that is, an express appreciation that the local is inherent to aggregated clinical evidence. Zuiderent-Jerak (2007), and subsequently Engel and Zeiss (2014), building on Timmermans and Berg's (1997) notion of 'local universalities', have thus referred to the need for 'situated standardisation' in the development of clinical guidelines: "Situated standardisation means that standards are practised in a situated manner, by assessing what the role of the guideline is in a particular service delivery situation and then adapting it respectively" (Engel and Zeiss, 2014: 205). As elaborated by Engel and Zeiss (2014) in relation to MDR-TB guidelines, this allows for local innovation within the confines of control, such as healthcare staff 'going beyond' what is officially sanctioned by the guidelines. Within the context of a clinical trial, this kind of local adaptation may well result in a so-called 'protocol deviation'.

Is there a place for 'situated standardisation' in multicentre trials? Some would argue that it already exists, and indeed recent methodological developments, namely so-called adaptive trial designs, suggest that there is a growing acceptance of adaptation over rigid standardisation (Montgomery, 2017). It is not only in biostatistics and trial methodology that such developments are occurring; in the critical medical humanities, Savransky and Rosengarten (2016) recently offered a different take on the ontology of health and disease, regarding such processes as always situated achievements:

> [W]hile RCTs locate mechanisms through the abstraction of discrete, isolated entities and variables from their 'confounding' environments, the practice of situating, by contrast, does not allow such clear-cut distinctions. Rather, both an object and its situation are entangled, spatially and temporally, to one another such that both become co-determined through their specific, reciprocal transactions and exchanges. (Savransky and Rosengarten, 2016: 170)

While their proposition provokes a series of questions about how to produce evidence of effect that can be generalised across settings and situations, it remains – as they themselves acknowledge – 'a fiction' (Savransky and Rosengarten, 2016: 171). As I have shown in this paper, there is an element of fiction to the 'real world' settings of the pragmatic trial, but that does not mean we should discount such evidence out of hand. Rather, what is needed are critical analyses of how the gap between evidence and implementation is forged, and a dialogue between the biomedical sciences and the critical social sciences as to better ways forward.

Conclusion

In this paper, I have sought to move beyond the dualism of critique and complicity which frames this special issue, by highlighting the ambiguity of standards in practice in the context of a global health trial. Sociological studies of standard-setting emphasise that standards and standardisation are not inherently good or bad, and demonstrate the ways in which standards can be made to work in local situations (Timmermans and Epstein, 2010). That is, standards tend not to be rigidly adopted in practice, but are most successful when they incorporate a degree of – but not too much – flexibility (Lampland and Star, 2009; Engel and Zeiss, 2014). The drive towards 'sensible' clinical trials, which has to date focused

on standardisation in the regulatory realm, would benefit from a broader appraisal of the forms of social control which suffuse the experimental process. The introduction of standards into the material environment transforms the existing social order, or 'social software', of clinical trial sites. We therefore need more thoughtful consideration of how the proliferation of standards accompanying increasing levels of clinical trial traffic in some places not only erases what is 'representative' about these places but also troubles what is meant by 'capacity building'. I have suggested in this paper that a productive way forwards is to propagate an appreciation of the fact that clinical trials need not only to be 'sensible' but also 'situated'.

Acknowledgements

This work would not have been possible without the trust, curiosity and enthusiastic participation of colleagues at the clinical trials unit in the UK who provided ethnographic access for this research. I would like to thank those working on the trial in question in the UK, South Africa and Vietnam for allowing me to be with them for the duration of the ethnography, and for giving up their time to be interviewed for this research. For insightful and constructive feedback on earlier drafts of the paper, thanks are due to Javier Lezaun, Salla Sariola, Patricia Kingori, Nora Engel, and two anonymous reviewers. The research leading to these results received funding from the European Research Council under the European Community's Seventh Framework Programme (FPT/2007-2013) / ERC grant agreement no. 263447 (BioProperty). Analysis and writing was supported by a VENI grant from the Netherlands Organisation for Scientific Research (NWO), grant no. 451-11-003.

References

Armitage J, Souhami R, Friedman L et al. (2008) The impact of privacy and confidentiality laws on the conduct of clinical trials. *Clinical trials* 5(1): 70-74.

Baigent C, Harrell FE, Buyse M, Emberson JR and Altman DG (2008) Ensuring trial validity by data quality assurance and diversification of monitoring methods. *Clinical trials* 5(1): 49-55.

Berg M (1997) *Rationalizing Medical Work: Decision Support Techniques and Medical Practices*. Cambridge, MA: MIT Press.

Bowker GC and Star SL (1999) *Sorting Things Out - Classification and its Consequences*. Cambridge, MA: MIT Press.

Cambrosio A, Keating P, Schlich T and Weisz G (2006) Regulatory Objectivity and the Generation and Management of Evidence in Medicine. *Social Science & Medicine* 63(1): 189-199.

Castel P (2009) What's Behind a Guideline? Authority, Competition and Collaboration in the French Oncology Sector. *Social Studies of Science* 39(5): 743-764.

Charmaz K (2006) *Constructing Grounded Theory*. London: Sage.

Cook D, Moore-Cox A, Xavier D, Lauzier F and Roberts I (2008) Randomized Trials in Vulnerable Populations. *Clinical trials* 5(1): 61-69.

Cousins S (2016) Funding shortfalls put target to end TB epidemic by 2035 at risk. BMJ: *British Medical Journal (Online),* 355.

Davis C and Abraham J (2013) *Unhealthy Pharmaceutical Regulation: Innovation, Politics, and Promissory Science*. New York: Palgrave Macmillan.

Engel N and Zeiss R (2014) Situating Standards in Practices: Multidrug-resistant Tuberculosis Treatment in India. *Science as Culture* 23(2): 201-225.

Epstein S (2007) *Inclusion: The Politics of Difference in Medical Research*. Chicago & London: University of Chicago Press.

Fisher JA (2009) *Medical Research for Hire: the Political Economy of Pharmaceutical Clinical Trials*. New Brunswick: Rutgers University Press.

Frick M (2016) Report on tuberculosis research funding trends, 2005-15. No time to lose. Treatment Action Group.

Fujimura J (1992) Crafting science: standardized packages, boundary objects, and "translation" In: Pickering A (ed) *Science as Practice and Culture*. Chicago: Chicago University Press, pp. 168–211.

Granger CB, Vogel V, Cummings SR, Held P, Fiedorek F, Lawrence M, et al. (2008) Do we need to adjudicate major clinical events? *Clinical trials* 5(1): 56-60.

Groves T (2009) Senior UK researchers campaign for sensible guidelines on regulating trials. *BMJ: British Medical Journal* 339: b3671.

Jordan K and Lynch M (1998) The dissemination, standardization and routinization of a molecular biological technique. *Social Studies of Science* 28: 773-800.

Knaapen L (2013) Being 'evidence-based' in the absence of evidence: The management of nonevidence in guideline development. *Social Studies of Science* 43(5): 681-706.

Knaapen L, Cazeneuve H, Cambrosio A, Castel P and Fervers B (2010) Pragmatic evidence and textual arrangements: A case study of French clinical cancer guidelines. *Social Science & Medicine* 71(4): 685-692.

Lampland M and Star SL (2009) *Standards and Their Stories: How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life*. Ithaca, NY: Cornell University Press.

Lang T, Cheah PY and White NJ (2011) Clinical research: time for sensible global guidelines. *Lancet* 377(9777): 1553-1555.

Latour B (1983) Give me a laboratory and I will raise the world. In: Knorr Cetina K and Mulkay M (eds) *Science Observed: Perspectives on the Social Study of Science.* London: Sage, pp. 140-170.

Light DW and Lexchin JR (2012) Pharmaceutical research and development: what do we get for all that money? *BMJ: British Medical Journal* 345.

Marks HM (1997) *The progress of experiment: science and therapeutic reform in the United States, 1900-1990.* Cambridge: Cambridge University Press.

Martin P, Abraham J, Davis C and Kraft A (2006) Understanding the "Productivity Crisis" in the Pharmaceutical Industry: Over-regulation or Lack of Innovation? In: Webster A (ed) *New technologies in health care: Challenge, change and innovation.* Palgrave Macmillan, pp. 177-193.

Montgomery CM (2017) From Standardization to Adaptation: Clinical Trials and the Moral Economy of Anticipation. *Science as Culture* 26(2): 232-254.

Müller-Rockstroh B (2012) Appropriate and Appropriated Technology: Lessons Learned from Ultrasound in Tanzania. *Medical Anthropology* 31(3): 196-212.

Petryna A (2009) *When Experiments Travel: Clinical Trials and the Global Search for Human Subjects.* Princeton: Princeton University Press.

Petty J and Heimer CA (2011) Extending the rails: How research reshapes clinics. *Social Studies of Science* 41(3): 337–360.

Savransky M and Rosengarten M (2016) What is nature capable of? Evidence, ontology and speculative medical humanities. *Medical Humanities* 42(3): 166-172.

Shurlock B (2013) Sensible guidelines for clinical trials: are current European regulations 'a gift to America'? *European Heart Journal* 34: 1767–1771.

Sunder Rajan K (2007) Experimental values: Indian clinical trials and surplus health. *New Left Review* 45: 67-88.

Thorpe KE, Zwarensteinb M, Oxmanc AD et al. (2009) A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *Journal of Clinical Epidemiology* 62: 464-475.

Timmermanns S and Berg M (2003) *The gold standard. The challenge of evidence-based medicine and standardization in health care.* Philadelphia: Temple University Press.

Timmermans S and Berg M (1997) Standardization in Action: Achieving Local Universality through Medical Protocols. *Social Studies of Science* 27(2): 273-305.

Timmermans S and Epstein S (2010) A World of Standards but not a Standard World: Toward a Sociology of Standards and Standardization. *Annual Review of Sociology* 36: 69-89.

Yusuf S, Bosch J, Devereaux P and al e (2008) Sensible guidelines for the conduct of large randomized trials. *Clinical Trials* 5: 238-239.

Zuiderent-Jerak T (2007) Preventing Implementation: Exploring Interventions with Standardization in Healthcare. *Science as Culture* 16(3): 311-329.

Zuidgeest MGP, Goetz I, Groenwold RHH, Irving E, van Thiel GJMW and Grobbee DE (2017) Pragmatic trials and real-world evidence—an introduction to the series. *Journal of Clinical Epidemiology*.