

Valuing height: diagnosis, valuation and the case of idiopathic short stature

Michael Morrison

Nuffield Department of Population Health, Centre for Health, Law and Emerging Technologies (HeLEX), University of Oxford, Oxford, UK 4

Abstract This paper proposes a 'valuographic' approach to diagnosis, exploring how values and valuation practices are implicated in the contested diagnostic category of idiopathic short stature (ISS). ISS describes children who are 'abnormally' short but do not have any other detectable pathology. In the USA growth-promoting hormone therapy has been approved for ISS children, since 2003. However, no other jurisdiction has approved this treatment and the value of ISS as a diagnostic category remains disputed among healthcare professionals. Drawing on qualitative interviews with paediatric endocrinologists in the UK and the US, this study presents a historical snapshot illustrating how the problematisation of ISS as a diagnosis involved multiple registers of value including epistemic, economic and moral calculations of worth. Contestation of the diagnosis was not just about what counts but about what ought to be counted, as respondents' accounts of ISS gave differential weight to a range of types of evidence and methods of assessment. Ultimately what was at stake was not just the value of increased height for short patients, but what it meant to properly practice paediatric endocrinology. Consideration is then given to how a valuographic approach can be applied to sociological studies of diagnosis more broadly.

Keywords: Contested diseases, Discourse, STS (Science and Technology Studies), Medical practice/medical work, Children, Interviewing (qualitative)

Introduction

In their call for an explicit sociological focus on diagnosis, Jutel and Nettleton (2011) argued that not only was diagnosis 'central to the practice of medicine' (Blaxter 1978: 9), but that its importance was also recognised in a wide variety of sociological analyses of medicine, health and illness, albeit often overshadowed by theoretical frameworks such as medicalisation. They marshalled this body of work to support their claim that diagnosis, considered as the classifications of disease, the process of assigning individuals to those categories and the

consequences of such classificatory activity, is a thoroughly relational, social and value-laden affair that warranted further sociological attention (see also Jutel 2009).

While work in the sociology of diagnosis has arguably restored 'diagnosis' to conceptual prominence, the treatment of how value and values impact diagnosis is somewhat out of step with developments elsewhere in sociology. Both value (especially economic value) and values (understood as cultural and moral registers of worth), are often treated as explanations for particular outcomes or behaviours. However, scholars working at the intersection of economic sociology and Science and Technology Studies (STS) are increasingly treating matters of value as something to be explained as part of a critical sociological analysis (Lamont 2012, Muniesa 2012). The recent edited collection 'Value practices in the life sciences and medicine' (Dussage et al. 2015a) builds on this work to set out the key tenets in what the authors describe as a 'valuographic' approach to the life sciences and medicine.

'Value' in a valuographic analysis can be understood in the sense of 'worth'; it encompasses both economic and other quantified calculations of what entities are worth (usually considered as 'value') and moral, cultural and epistemic judgements of what is 'worthwhile' or 'worth doing' (commonly regarded as 'values'). This does not imply that any one of these categories is reducible to another, e.g. that 'ethics has really been economic all along'. Rather, multiple registers or 'orders of worth' are recognised (Lamont 2012).

Valuography also adopts a pragmatic approach to worth. Instead of treating 'value' as either intrinsic to objects or a purely subjective designation, a pragmatic stance regards 'value' as the outcome of processes of valuation (Muniesa 2012). 'Processes' is broadly interpreted to include a variety of practices, tools, organisational arrangements and devices through which value is assigned to people, objects, goals and other phenomena. The idea of worth as something made in practice, and thus amenable to empirical study, is easier to grasp in relation to quantitative forms of value. It is not difficult to imagine the price of goods, for example, as the outcome of calculations balancing cost of production and consumer demand. However, valuation also encompasses qualitative assessments; the aesthetic merit of a painting, the quality of a wine, or the fairness of a system of organ transplantation. Here, it is helpful to consider that while 'value ideals' such as artistic merit, fairness or justice, and

cultural, ethical and professional norms (e.g. that physicians should 'first do no harm') generally pre-exist any given situation, how those ideals and norms come to be interpreted and implemented in any particular situation will depend on the practices through which they are operationalised (Roscoe 2015).

The purpose of this paper is to illustrate how a valuographic approach could contribute to sociological studies of diagnosis. Since valuography is explicitly attuned to empirical study of value-making in, and through, practice, it is appropriate to take a case study approach. Accordingly, I will first consider some general features of how valuography can be integrated into existing sociological perspectives on diagnosis before illustrating how it might be applied to the specific diagnostic category of idiopathic short stature (ISS).

'Idiopathic short stature' means 'short stature for which there is no known cause' (c.f. Beaney 2013). It is a diagnostic category applied to children who are considered to be 'abnormally' short, where abnormal (short or tall) stature is defined statistically as being at least 2 standard deviations (SDs) above or below the mean population height adjusted for age and sex. The notion of a specific diagnostic category to designate children who are significantly short, but otherwise in apparent physical health came to prominence in 2003 when the United States Food and Drug Administration (FDA) granted regulatory approval for the pharmaceutical human growth hormone (hGH) to treat children with ISS (Conrad and Potter 2004).

A number of features recommend ISS as a case study for valuography of diagnosis. Dussage et al. (2015a: 272) recommend matters of controversy and contestation as fruitful sites for study because 'conflicts evoke rich articulations about the values at stake' while also being instances that illustrate the mutability of values, as what counts as worthwhile or worth doing often changes over the course of a controversy. The idea that 'normal short' ISS children might be in need of growth-promoting treatment with an expensive prescription-only biopharmaceutical was controversial before and after the FDA decision of 2003. The merits of ISS as a diagnostic category, the value of giving hGH to children with this diagnosis and the rationale for medical intervention itself were all subject to considerable, occasionally

acrimonious, debate in the academic medical literature (Allen and Frost 1990, Sandberg and Colsman 2005).

Writing about the wake of the Chernobyl and Fukushima disasters, Dussauge et al. argue:

Valuations of life, knowledge and money become matters of concern. Whose assessments of radioactive fallout are valid? Whose assessments of lives, quality of life, and livestock? Assessments of different values are intertwined: economic value, risk values, health values, quality of life (Dussauge et al. 2015b: 11).

Similar issues can be observed in the disputes over ISS. What is at stake for healthcare professionals includes both the value of ISS and the validity of treating that patient population. Are the statistical calculations upon which a diagnosis of ISS is based adequate to distinguish between health and disease? Is the most appropriate way of calculating the value of hormone therapy for ISS children the amount of additional height gained per year of treatment, the economic cost per patient, or something else? What makes short stature sufficiently undesirable to warrant medical intervention in the first place? Is it a physical disadvantage, a barrier to psychological wellbeing, or a matter of social prejudice? What justifications were employed and how was worth attributed or disputed in these debates?

Towards a valuographic approach to diagnosis

The very process of making a diagnosis can be understood as an evaluative practice. The addition and subtraction over time of particular categories or 'cells' in the overall matrix of diagnostic classifications 'provides a cultural expression of what society is prepared to accept as normal and what it feels should be treated' (Jutel 2009: 279). This is replicated at the level of individuals; through diagnosis the patient's narrative of illness is either legitimated (given worth) as disease or rejected (devalued) as delusional or deviant behaviour. Being diagnosed also has consequences for how individuals are valued by society; acting as a gateway to care and social support, but potentially also bringing stigma and social abjection.

Valuography not only recognises multiple registers of worth, but also it allows that in 'real world' situations, different registers are often entangled in making calculations of worth. We

can see this when, for example, assessments of the lived experience of illness inform economic calculations of value through quality-adjusted life-years. Diagnosis too, has been characterised as a site where different types of concern become entangled, representing 'a focal point where numerous interests, anxieties, values, knowledges, practices and other factors merge and converge' (Jutel and Nettleton 2011: 798). The multiple measurements, instrumental readings and registers of assessments that make up an overall diagnostic process are themselves subject to valuations. Rather than their being a clear, pre-determined value hierarchy of evidence for diagnosis, studies of clinical work especially those informed by STS, have demonstrated that making a diagnosis involves weighing up different kinds of evidence, from the patients' account of their illness to instrumental measurements, laboratory results, clinical guidelines, and personal experience and then making a (clinical) judgement about how to proceed (Latimer et al. 2006).

Acts of ranking, comparison and classification like these are all identified by Lamont (2012) as sub-processes of valuation. They are not simply about what is counted and known, but they involve judgements about what should count and what is worth knowing. Further, diagnosis is rarely a disinterested calculation; it is often made in the context of guiding what the physician can do with, and for, the patient in front of them (Mol and Elsman 1996). Diagnostic judgements thus also have the potential to segue into judgements about prognosis, treatment and wellbeing (Lowy 2015). Each of these elements, as will become evident, is relevant to understanding the case of ISS.

Case study: the history and context of ISS

To understand how childhood short stature first came under medical management it is necessary to look back to 19th century studies of growth and height when public health officials began to routinize the application of statistical measures to assess the heights of adults and children (Armstrong 1995, Tanner 1981). The inclusion of children in these surveys was prompted by concern for their adequate growth and development in an age of industrialisation, widespread poverty and child labour. In addition, a range of conditions including (then untreatable) diabetes, rickets and achondroplasia were all associated with retarded childhood growth and short stature. Statistically abnormal height was not initially considered to be a disease in itself, but was regarded as a symptomatic marker of potential

underlying pathology which warranted close medical attention (Armstrong 1995). The main device for separating 'short stature' from 'normal stature' was, and remains the growth chart. Growth charts are based on the statistical concept of the normal distribution curve, in which the distribution of height values, for a given age and gender, are arranged in a bell-shaped curve with the mean height as the central point.

The measurements of any given child are plotted on the chart and compared against the mean value to see how far above or below average height they are at any given age. Applying statistical calculations initially developed to assess the likelihood of errors in astronomical readings, 'normal' values are considered to lie within 2 SDs of the mean value while readings more than 2 SD are considered 'abnormal' (Tanner 1981). As with many formal tools of assessment the growth chart is presented as an objective measure of 'what is', such that the value judgements that inform its design are effectively hidden. From a valuographic perspective we can see how this comparison with the mean produces shortness as a deficit, giving the statistical average normative value as the standard for what is appropriate, desirable and healthy. When hGH was first isolated in the late 1950s there was already a longstanding medical tradition of monitoring height as an index of childhood wellbeing and an association between short stature and poor health. The new growth hormone was initially tested in children with a wide range of known or assumed causes of short stature, and some with no clear diagnoses. This was partly a deliberate strategy to investigate the hormone's role in biological growth, and partly because some clinicians hoped the new treatment might be a universal remedy for childhood short stature (Morrison 2015).

Those early trials found that hGH did not produce a growth response in all short children. Supply of the hormone was severely limited due to the fact it could only be extracted from human pituitary glands collected at autopsy. It was therefore reserved for children who were the most deficient in height and had lowest levels of endogenous growth hormone in their blood. This formed the new diagnostic classification of growth hormone deficient (GHD) short stature. In the USA, a small amount of pituitary hGH was also available for academic paediatric endocrinologists conducting experimental research on growth and growth disorders. A common theme in this research was examining whether hGH could have a growth promoting effect in other classifications of short stature beyond GHD. As causes of short stature became

gradually subdivided into those that did not respond to hGH, such as achondroplasia, and those that demonstrated some improved growth rates such as Turner and Prader–Willi syndromes (both genetic conditions), a ‘remainder’ category emerged – children who were short but with no clear underlying pathology. Initially called ‘normal variant short stature’ this was later renamed ‘ISS’. It was perhaps inevitable that some academic paediatric endocrinologists then began to explore the effect of hGH on ISS children, although this was strictly categorised as medical research not therapy at the time.

In 1985, pituitary hGH was replaced by a biosynthetic form of the drug made by Genentech (now an independent subsidiary of F. Hoffman La Roche) and Eli Lilly. With supply no longer limited by the availability of pituitaries, a series of trials were conducted in categories of short stature known to respond to the pituitary form of the hormone. Synthetic hGH was first approved for GHD and in subsequent years for Turner syndrome, Prader–Willi syndrome and small-for-gestational age children (Author 2015). Growth hormone was approved in each of these categories by the FDA and the European Medicines Agency. The exception was ISS where only the US approval was ever sought or granted. This does not mean that all ISS children in the USA could get hGH and no ISS children in other jurisdictions like the UK were treated. Selected ISS children in the UK were, and are, treated with hGH through the discretionary off-label prescriptions, while only some insurance plans in the USA actually include ISS so access is still limited by ability to pay.

In this context, the close connection between ISS as a diagnosis and hGH as a treatment makes sense. The category of ISS owes its existence to the systematic study of hGH as a growth promoting agent. It remains controversial because it is the only one of the several indications for hGH defined by short stature alone,¹ but it also represents the apogee of an expansion in hormone treatment for short stature rather than an ‘enhancement’ use appearing from nowhere. Indeed, it has been argued elsewhere (Author 2015) that the presence of a biological component (a pathological lesion) in the other diagnostic categories in which hGH is applied deflected serious scrutiny of the rationale for treating short stature per se, and that part of the controversy over ISS is because it forces attention to the question of what is actually gained by making any short child a little less short.

Methods

The empirical data analysed in this paper derives from a doctoral research project comparing use of hGH in the UK and the US across a range of accepted and contested applications including ISS (Author 2008). Between May and December 2006 16 semi-structured interviews were conducted: six with UK-based endocrinologists; nine with clinicians (eight endocrinologists and one paediatric psychiatrist) based in North America, and one supplementary interview with a representative from a patient charity for childhood short stature. Interviews were conducted face-to-face and by telephone, recorded, transcribed and analysed by the author. Informed consent was obtained from all participants. To protect participants' professional standing and encourage open discussion, interview transcripts were anonymised and participants' assigned a code based on their location e.g. 'NAM3' to signify 'North American interviewee number 3'.

Interviews took place within a few years of the controversial 2003 FDA decision on ISS and disputes over the diagnosis were still a matter of professional concern in paediatric endocrinology. ISS was not the sole focus of that project and the sample included adult endocrinologists who were not asked about ISS. Only 10 respondents had sufficient experience of paediatric practice to discuss ISS: four UK-based paediatric endocrinologists, four US-based paediatric endocrinologists, one US-based paediatric psychiatrist and one paediatric endocrinologist based in Canada. All ten had been authors on at least one published academic paper about the condition and had experience of assessing ISS patients in the clinic. However, some interviewees were more engaged with the dispute than others. This is reflected in the interview data analysed here where some respondents have greater prominence than others. Those less quoted did not offer a different perspective, but expressed similar points albeit sometimes less fulsomely. Further information about respondents is provided in Table S1.

The small sample size and age of the data place several limitations on the scope of analysis. Although the data was adequate for the original study, the subset on ISS cannot be said to have reached theoretical saturation. It is likely further interviews would provide additional perspectives on ISS and the findings presented here cannot be considered representative of endocrinologists in the UK, the US or Canada. Revisiting old data means the current analysis

necessarily represents a historical 'snap-shot' in which the opinions and practices reported in relation to ISS, may not reflect the current state of endocrinology. However, the primary purpose for revisiting this data is not to provide a comprehensive analysis of diagnostic decision making in ISS but to use ISS as a case study to illustrate how a valuographic approach and perspective can usefully be incorporated into a sociological analysis of diagnosis. This does not negate the limitations of the data but it does make them less critical when considering the task at hand.

The original study did not employ an explicitly valuographic approach. Interviews were designed to elicit how respondents' justified, contested or problematised use of hGH in a range of patient populations. For the present analysis a subset of the original data, in which respondents discussed ISS, has been reanalysed to draw out specific value-related aspects. This incorporated inductive and deductive aspects, since interview data had already been given some structure through the original questions and by the commitment to look for aspects of discourse relating to values, conduct, weighting of evidence and diagnostic decision-making, whilst at the same time remaining open to novel patterns and framings emerging during analysis. As with the original study, the account produced from (re)analysing these interview transcripts does not purport to be a recreation of what 'actually happens' in the endocrine clinic. The interviews do reveal how respondents' give value to certain tools and types of evidence and devalue others, how they justify particular courses of action and what they present as proper or improper conduct. These were 'elite' interviewees; the majority of respondents were high status academics, well respected in their fields and used to imparting information and explanation to junior colleagues and students (i.e. persons of a similar status and age to the author at the time of the original study) making it more likely that they would draw on familiar norms and value-registers during the interviews.

Results

Hierarchies of evidence in diagnostic decision-making

The category of 'ISS' depends on the use of the growth chart as a device to produce, and thereby make distinguishable, normal and abnormal short stature. Statistical assessment of height-for-age and growth rate are common to most, if not all, diagnostic categories that

include short stature as a symptom, including GHD short stature, Turner syndrome and others. However, both North American and British paediatric endocrinologists problematised the idea, which underpins the category of ISS, that these measurements alone were a sufficient basis for making a diagnosis:

It's not a diagnosis, it just describes a collection of small children, [. . .] some of them may be normal small children, others may have specific abnormalities (UK1).

I think it is clearly a heterogeneous group of patients and it is going to encompass children who are normal and are just at one end of the bell shaped curve, and [. . .] it's going to include children who have significant pathology (NAM5).

For these respondents, ISS failed as a satisfactory diagnostic category because it failed as a dividing practice. The measurement of height alone was not considered sufficient to reliably separate healthy children from those with 'significant pathology'. This reflects the provisionality inherent in the 'idiopathic' label. As Beaney (2013: 128 emphasis added) observes 'an idiopathic disease is so-called because we have not discovered a cause but expect one to exist'. Downplaying the epistemic value of height as an adequate basis for diagnostic judgement opened a rhetorical space where respondents could emphasise the desirability (i.e. the greater worth) of alternative criteria that might resolve the diagnostic uncertainty of 'idiopathic' short children

[Y]ou can't consider them as a single group, you have to look within that general category at the child who may really respond very well and if you look hard you can find them [. . .] I think it doesn't help particularly treating children of short parents. (NAM1).

I do not exclude the possibility that there's some very short kids who may be in the familial/ idiopathic category who may benefit from growth hormone, the problem is we can't identify those kids easily and all these manipulations of looking at growth hormone receptor and growth hormone releasing hormone (GHRH) have not panned out yet (NAM8).

The most commonly discussed type of alternative diagnostic criteria was various molecules (insulin-like growth factor 1, GHRH, cell surface receptors). This offers an example of the way the value of different types of evidence (physical measurements, biological measurements) was ordered into hierarchies in processes of diagnosis. In diagnosing short stature, physical-

statistical measurements were relegated to the status of 'symptoms' while the power to identify a 'causal' marker of pathology was attributed to signs of disease located 'deeper' within the body – in organs, cells and molecules (c.f. Aronowitz 2001).

This stance was echoed in the clinical literature of the time where it was hoped further laboratory investigation of the biology of short stature would:

chip away at the monolith of ISS and, by identifying new pathogenic mechanisms, [. . .]
gradually reduce the number of children who currently fall into this rather
unsatisfactory category (Blair and Savage 2002: 329).

There are similarities between this appeal to further research and the deferral of judgement in diagnosis of children with dysmorphic syndromes reported by Latimer et al. (2006). In both cases deferral is a response to the combination of uncertainty and anticipation where the moment of definitive evaluation – finding the cause that is 'expected to exist' – is projected into an imagined 'future in which classification becomes sedimented in a technoscience that will continue to press back the boundaries of the [. . .] unknown' (Latimer et al. 2006: 622). In the dysmorphia clinic, deferral sustained medical authority by allowing children with suspected dysmorphia to remain under the supervision of the clinic, awaiting further technical assessment and the clarity it was anticipated to bring (Latimer et al. 2006). For paediatric endocrinologists dealing with ISS children, molecular analyses had 'not panned out yet' (NAM8) and remained a more distant, less accessible future. Instead, respondents suggested other ways in which they could sustain medical authority, and discharge their responsibility to the patient in front of them, such as the suggestion, contained in the quotes above, of distinguishing between cases of ISS based on which subsets of children might benefit from treatment.

Normative registers in clinical judgement

'Benefit' here meant respond to growth hormone treatment (by growing faster and attaining a taller final height). The idea that a bio-physical response to treatment ought to be a criterion for receiving hGH went back to the experimental studies of the twentieth century when various categories of short statured children were given hGH and classified as growth hormone responsive (and eligible for treatment) or growth hormone insensitive (and

ineligible for further doses of the drug) depending on whether or not they demonstrated a growth response. This rationale carried its own normative implications. It reflected an underlying value judgement that that being taller is better for children and, by inference that being short was a disadvantage. By this logic, short children who could respond to hGH were entitled to treatment with hGH. As with other types of short stature, response to treatment was measured in terms of changes in childhood growth rate over time and, ultimately, in the height of treated children compared to the heights predicted for the same children if they remained untreated. The inherent unknowability of how any given child would have grown without treatment 'blurs the boundary between diagnosis (through-knowing) and prognosis (before-knowing)' (Lowy 2015: 195). Instead other criteria – such as the experience-based judgement that 'it doesn't help particularly treating children of short parents' proffered by NAM1, came to have weight in making judgements about which ISS children ought to get hormone therapy. These prognostic criteria (which children might respond well) fed back into diagnostic decision-making.

Whether or not children came from a 'short' family was not part of the official definition of ISS, but it could be utilised in decisions about treatment eligibility through the endocrinologists' clinical judgement. This demonstrates how making a diagnostic assessment can involve multiple normative registers; that is multiple ideas of what is worth doing and what ought to be done. As well as looking to reliably distinguish healthy from sick children, assessments of ISS children were also being made with a view to identifying which children might produce a response to treatment and which might not. This shift is subtle, because they are often discussed together as in the quotes above, but it is significant. It not only changes the focus of evaluation, but the nature and scope of calculations involved. As one register of worth was displaced by another, new logics were introduced as well as new sites of contestation, concerning what kinds of evidence are needed and how they should be valued, interpreted and acted upon. While changes in height and growth rate are readily quantified, the value of these changes had proven much more difficult to assess. Economic assessment of hGH treatment in ISS is a case in point. Consider the following perspectives representing opposite ends of the spectrum:

What I object to in the literature for the last 20 years is the fact that it's a 'limitless' or 'unlimited' supply of growth hormone. That's not true; someone has to pay for it (NAM8).

I do take money out of the equation, because if you just concentrate on how much it costs per inch or centimetre it doesn't sound as good, but if you're trying to look at the whole picture . . . So [] my bias is to treat (NAM5).

These were not simply disputes about appropriate weighting of empirical evidence, they were normative disputes about whether economic cost should or should not be a factor in calculating whether intervention with hGH was worthwhile. The contestation revealed differing ideas about what it meant to responsibly practice paediatric endocrinology. Both respondents offered justifications of their positions by appealing to ideas of appropriate professional conduct (c.f. Nordlund 2015):

I think one can't ignore the ethics and the economic arguments, so because it is an expensive product (NAM8).

I don't deal with the economics I make decisions based on trying to be the best doctor I can for that child, that family and so obviously I have a limited practice (NAM5).

In the first quote, the locus of the physicians' responsibility was shared between the patient and those responsible for paying for the treatment (from a constrained budget). The magnitude of the cost of hGH meant the 'responsible' physician was impelled to give economic factors weight in calculating the value of intervention. By contrast the second respondent posited the physicians' primary responsibility as being to 'that child, that family' and justified ignoring the financial impact of providing therapy precisely because money distorts or distracts from the primacy of that relationship. These appeals to contrasting ideas of 'good professional conduct' and what should count as an appropriate method, relevant evidence, benefit and harm, were echoed in another part of the ISS debate. Besides cost, another point of contestation was arguably even more fundamental to the whole issue of the medical management of short children. If the treatment for (idiopathic) short stature is growth promoting treatment, what is it about shortness that warranted intervention in the first place?

The 'problem' of short stature

The science writer Stephen Hall has observed that:

[t]he notion of shortness as a psychological disadvantage – indeed, disability, – runs deep and persistently through a huge scientific literature on human physical stature (Hall 2006: 16).

Similarly, in an article in *Pediatrics* in 1990, Allen and Frost reviewed a range of studies suggesting that short stature can be detrimental to an individual's development, psychological wellbeing, social and financial success and concluded that: 'discrimination based on height- "heightism"- pervades American life' (Allen and Frost 1990: 17). In earlier parts of this section we saw how ISS had been framed in physical/statistical terms as a deficit in growth rate and height and in biochemical terms, as a site of (future) classification based on molecular or genetic markers. Here a third framing – a third ontology of short stature – entered the debate; shortness as a psycho-social burden. The suggestion that shortness by itself was a diagnosable handicap and a justification for growth promoting treatment had grown in prominence during the long debate over whether ISS children should get hGH. Respondents were certainly aware of the parameters of the discussion:

[I]f you're asking what is the rationale for treating short stature by itself . . . well I think there is a perception, you know rightly or wrongly and that's something that can be debated all day long, that children and adults who are significantly short are disadvantaged, that they are handicapped, that, you know, whether it has to do with social performance, or academic performance or marital opportunities or employment opportunities, and you know you can find literature on both sides of the fence on that issue, none of the data are perfect or are very good, but certainly there's a perception in families that a small child is disadvantaged (NAM 5).

In this relatively brief answer a considerable range of different types of problems or deficiencies associated with being short were posited, from educational attainment to psychological wellbeing. There is also considerable uncertainty and indeterminacy; the disadvantages of shortness 'can be debated all day long'. It is not clear whether disadvantage (if any) accrued to the child in the present or to the (short) adult they were expected to become. This echoes findings from the sociology of childhood where it has been argued that

‘adult’ social orders (like medical classification) are challenged by the inherent ambiguity of children as both individuals in their own right and ‘not-yet-adults’ (Lee 1999). As with uncertainty occasioned by endocrinologists’ rejection of height as a definitive diagnostic criterion, ambiguity opened up space for new and competing terms of assessment, and ideas of what was at stake and what should be done, to be articulated and justified.

In this case, different ideas of how a psycho-social burden of short stature might be understood and framed, or what Dussage et al. (2015a) describe as different ‘ontonorms’, mobilised different methods for measuring the distress caused by shortness and different criteria for evaluating the evidence produced. Formal testing for quality of life (QoL) and psychological wellbeing was one way of framing the putative psycho-social burden of shortness. This required additional expertise, commonly from paediatric psychologists, and posited any psycho-social deficit as a collective problem, detectable in the population of short children (by comparison with a control population of normal-statured children) and quantifiable through the use of standardised QoL questionnaires. In this logic the appropriate format for generating evidence of psychological harm was the clinical trial, a number of which had been carried out in both the UK and USA:

[W]e devised a trial to see whether we could improve the adverse psychological factors that we could identify in short children, well when we got hold of a bunch of short children we couldn’t find anything to test (UK4).

[T]he psychosocial for the ISS NIH study was zero, they didn’t show any gain, they didn’t show any loss, they also didn’t show there was anything wrong with them when they started, so . . . Duh! (NAM8).

The failure of these methods to generate evidence of psychosocial suffering caused by being short was invoked by some respondents to reject the idea of psychological distress as a valid problem conferring entitlement to growth-promoting therapy. In these discussions of the (ontological) validity of shortness as a cause of suffering and thus a potential source of entitlement to hGH treatment, the sense of diagnosis as a moral calculation of who deserves access to medical care was clearly evident.

Evidence, or its lack, from clinical trials using formal QoL or psychological testing did not settle the debate. Respondents who valorised the idea of short stature as a 'real' psychosocial burden, drew on an alternative conception of psychological harm as existing primarily for the individual child, with an accompanying method of analysis that did not rely on standardised tests or group analysis:

I think every patient [. . .] has to be individualised and so before I would say yes or no to any patient I, and I think every paediatric endocrinologist should do this, every doctor ought to sit down and learn about the child, does that child have confidence or is that a child who is very shy and retreats. Is the child bothered enough by being short that it's really affecting the mental development (NAM3).

Other US-based respondents echoed this emphasis on individual assessment of the child in the clinic: 'each endocrinologist has to pick and individualise it from child to child' (NAM5), and 'why the hell should we treat a disease, we should treat a child' (NAM2). The focus on individual need located the appropriate expertise for gauging distress, in terms of 'confidence' 'shyness' and 'mental development' (NAM3) in the clinical judgement of the paediatric endocrinologist. This acted as a counterpoint to the population-based data derived through QoL tests applied by psychologists in clinical trials.

Respondents favouring this line of argument noted that psychologists studying the psychosocial effects of short stature had their own disputes, the most controversial being over which was the correct population of short children to examine for psychological stress – those referred to clinics or in the general population. One US-based practitioner observed 'if you look at the data, there are two sets of data' (NAM2) and another noted that proponents of each approach 'reject each other's studies' (NAM8). This devalued QoL or psychological evidence as both 'contested' and outside paediatric endocrinologists' domain of expertise and thus not amenable to resolution by them. These approaches assigned different worth to particular kinds of evidence and expertise and both were again justified through contrasting accounts of what it meant to be a 'good' paediatric endocrinologist. The sample size in this study was too small to permit generalisation, but it can be seen throughout the results that US-based paediatric endocrinologists tended to minimise the relevance of financial constraints on prescription and promote individual-centred evaluation and intervention in ISS

children more than physicians based in the UK or Canada. It seems likely that these physicians' responses were shaped by the regimes of practices in which they are situated. However, this is not the same as saying that different healthcare systems simply determined the values at stake. There is nothing in the general market-orientation of the US healthcare system that directly devalues either QoL testing per se or evidence from clinical trials. Indeed many of the US paediatric endocrinologists interviewed would use both kinds of evidence in other clinical settings. This illustrates how existing value ideals, such as the value-orientation of a for profit health system, and responsibility to the individual patient as an ideal of good clinical care, intertwined and played out (i.e. were made in practice) in a specific way in a particular diagnostic dilemma. There was not one unanimous answer, even within a given professional community. Instead, several different accounts, emphasising different valuation practices were produced, yielding what in effect were contrasting models of the appropriate conduct of a paediatric endocrinologist and of medicine itself. In many ways, the 2003 FDA decision and the debates captured here proved to be a 'high water mark' for controversy about ISS. Debates about its legitimacy have somewhat faded from prominence without any clear resolution of the matter as paediatric endocrinologists in different countries reached various accommodations with the national and local healthcare settings in which they work. No regulatory approvals have been sought for ISS outside the US, and use of hGH in ISS children in other jurisdictions remains restricted to off-label prescriptions.

Discussion

Paediatric endocrinologists' discussions of ISS ranged across a number of interconnected themes; considering the diagnosis as category (debating its legitimacy as a diagnostic classification), process (what should be done with, or for, children assigned to that category) and consequence (how best to evaluate the costs and benefits of intervention). Rather than being regarded as a single process of evaluation and assessment, diagnosis can be characterised as involving multiple, intertwined sub-processes of valuation and (clinical) judgement. These processes involved calculations made in different registers of worth, and they were not discrete but entangled; moral evaluations were interlinked with economic and epistemic assessments. This demonstrates the utility of attending to the multiplicity of value/s and their interconnectedness through a valuographic approach.

Previous sociological accounts examined the expansion of patient populations receiving hGH, using theoretical frames of medicalisation (Conrad and Potter 2004) and pharmaceuticalisation (Author 2015). While these approaches have merit, they risk treating the diagnostic category of ISS itself as a mere facilitative device that permitted ever-greater sales of hGH. This obscures the extent to which contests over the meaning and legitimacy of ISS were central to the debates within the medical profession about which children should, or should not receive hGH. By contrast, this account shows how foregrounding diagnosis (c.f. Jutel and Nettleton 2011) and applying a valuographic lens (Dussage et al. 2015a) can complement and extend the picture provided by prior sociological studies. Moreover, in keeping with the main aim of this article, it suggests several ways in which valuographic approaches can benefit the sociology of diagnosis more generally.

We have seen that ISS sat at the intersection between different domains of medical practice, each with its own logics and value-orientations. The surveillance of childhood growth that created a pool of 'abnormal short' children for medical attention derives from public health and valued health at the population level, understood in terms of statistically normal development, as its desired outcome (Armstrong 1995). Biomedicine, the medicine of the clinic, anchors its epistemic authority in the detection of specific biological signs of disease located in the anatomical body of the patient; often granting greater epistemic worth, the 'deeper' within the body a sign can be detected (Aronowitz 2001). 'QoL' and psychological assessment are tools of the 'new public health', which emphasised risk and extended its gaze beyond the material body to encompass health behaviours, lifestyle factors and 'psychosocial' aspects of wellbeing (Armstrong et al. 2007).

ISS was rendered fragile as a workable diagnosis because the different kinds of evidence produced by these domains of assessment did not align in ways that could easily inform a decision about what to do next (Mol and Elsmann 1996). ISS children had statistically abnormal height, but lacked detectable biological basis for pathology, while the evidence for their psychosocial wellbeing was equivocal and contested. This opened up spaces for the exercise of clinical judgement and the creation of hierarchies of evidence (c.f. Latimer et al. 2006), embedded in normative claims about the value of different strategies for assessment and the

‘proper’ practice of medicine (c.f. Nordlund 2015). What was at issue was not only what was counted, but also what should count and what was worth doing: formal evaluation with psychological tests versus individual evaluation by a paediatric practitioner, the general population of short children or the particular child, the financial cost of treatment or the duty of care alone etc.

Bearing in mind the advice of Dussage et al. (2015a) that situations of dispute and conflict are useful sites for examining the making of values in practice, a valuographic approach could readily be extended to other cases where the proper boundaries of a diagnostic category or the processes of making a diagnosis are contested. Many of the tensions and faultlines discussed in relation to ISS are not unique to that case. For example, Lowy (2015: 199) described a similar tension between individual patient and population as the appropriate locus of care in her account of how prenatal diagnostic testing was shaped by contests between ‘gynaecologists and obstetricians, committed to ethics of individual centred medicine, and [. . .] public health experts and administrators who reason in terms of the cost/efficacy of health measures’. We could look for similar tensions and arguments about which kinds of evidence and conduct are considered legitimate and appropriate in other disputed conditions such as fibromyalgia or myalgic encephalomyelitis or in cases of expanding medical authority, many of which are already identified in the literature through labels of ‘disease mongering’, ‘medicalisation’, ‘pharmaceuticalisation’ and ‘human enhancement’. As with ISS, they invoke normative questions as to which accounts of suffering are valid, who deserves access to medical resources and under what circumstances, all of which are amenable to a valuographic perspective.

It is also important to acknowledge that important voices were missing from this account of ISS, most significantly those of patients and their families. Again, this suggests avenues for further exploration; our understanding of what is valued or devalued and how a judgement of worth is assigned through the diagnostic process could be significantly improved by supplementing interview data with ethnographic studies of clinical encounters between physicians and patients. Conflicts are not the only suitable sites for valuography. Spaces where diagnoses are revised and rendered unstable, and comparisons of how things are done differently across different spaces, also afford opportunities for seeing values in the making.

As Mol and Elsman (1996) observed, a given disease may be understood and assessed differently in different spaces, even within a particular hospital or care system. As the case of ISS illustrated, different tools used in diagnosis incorporate particular logics and values; for example the growth chart used to diagnose abnormal short stature was informed by a public health ethos that set the statistical average as the standard for what is desirable and healthy. However, the (changing) ways these tools were used in practice can also reconfigure those same values (Dussage et al. 2015a, Roscoe 2015). In at least some settings, the deficit of short stature produced by growth charts was (re)configured as a matter for individual patient-centred evaluation and treatment. More broadly, as the tools, devices, guidelines, protocols and standards used in defining and making diagnoses are revised and new ones introduced, values can be recalculated, some concerns will come to count while others may be displaced or disregarded.

Examining how the tools and devices that measure, rank and categorise the bodies of the diagnosed enact particular kinds of worth seems especially pressing in an era when technologies from genomic sequencing to artificial intelligence are heralded as 'revolutionising' diagnosis. Many new diagnostic technologies are implicated in the move to predictive and preventative medicine, based on identifying risk factors for potential disease rather than symptomatic illness. As with the shift from defining ISS to looking at which ISS children might respond to growth hormone, the transition 'from diagnosing people who have a disease to diagnosing people who might develop a disease fundamentally alters the calculus of benefits and harms involved' (Author 2016: 720 emphasis added). Valuography of diagnosis could usefully inform our understanding of the implications of an increasingly preventive, future-orientated medicine, for example in looking the setting of diagnostic thresholds and 'cut-off points' for conditions such as hypertension, considered by some to be 'overdiagnosed', or unpacking the values at play in debates over returning 'additional' findings uncovered during whole genome sequencing. Finally, diagnostic categories operate, in many spaces outside the clinic: in clinical trials and academic research programmes, in health technology assessment, in epidemiology, and in hospital budgeting and management. These are all spaces where diagnostic categories are involved in calculations of what is valuable and what is worth doing. Their consequences include economic impacts and organisational implications; from shaping what counts as an indication for drugs produced by

the pharmaceutical industry to national expenditures on public health. There are therefore likely to be many other cases where a valuographic approach could illuminate the interplay between economic and other registers of worth in defining, changing or sustaining diagnostic classifications and practices, and in decisions about who is and is not entitled to care.

Address for correspondence: Michael Morrison, Centre for Health, Law and Emerging Technologies (HeLEX), Nuffield Department of Population Health, University of Oxford, Ewert House, Ewert Place, Summertown, Banbury Road, Oxford OX2 7DD, UK. 5

E-mail: michael.morrison@dph.ox.ac.uk

Acknowledgements

I would like to extend my sincere thanks to the editors and the three anonymous reviewers whose supportive and insightful comments on previous versions of this manuscript have been invaluable. The original research utilised in this paper was funded by the Economic and Social Research Council (award number PTA-030-2004-00601).

Supporting information

Additional Supporting Information may be found in the online version of this article:

Table S1. Details of interview respondents discussing idiopathic short stature.

Note

1 The actual indication for ISS was set at 2.25 SD below the mean not the statistical cut-off of 2 SD. There is no explicit rationale for this extra increment but it may have been an attempt to reassure critics that the children being treated were ‘really’ short.

References

- Allen, D.B. and Frost, N.C. (1990) Growth hormone therapy for short stature: panacea or Pandora’s box?, *The Journal of Pediatrics*, 117, 16–21.
- Armstrong, D. (1995) The rise of surveillance medicine, *Sociology of Health & Illness*, 17, 393–404.

- Armstrong, D., Lilford, R., Ogden, J. and Wessely, S. (2007) Health-related quality of life and the transformation of symptoms, *Sociology of Health & Illness*, 29, 570–83.
- Aronowitz, R.A. (2001) When do symptoms become a disease?, *Annals of Internal Medicine*, 134, 803–8.
- Beaney, T. (2013) Classifying unknowns: the idiopathic problem, *Medical Humanities*, 39, 126–30.
- Blair, J.C. and Savage, M.O. (2002) The GH – IGF-1 axis in children with idiopathic short stature, *Trends in Endocrinology & Metabolism*, 13, 325–30.
- Blaxter, M. (1978) Diagnosis as category and process: the case of alcoholism, *Social Science & Medicine. Part A: Medical Psychology & Medical Sociology*, 12, 9–17.
- Conrad, P. and Potter, D. (2004) Human growth hormone and the temptations of biomedical enhancement, *Sociology of Health & Illness*, 26, 184–215.
- Dussage, I., Helgesson, C.-F. and Lee, F. (2015a) *Valuography: studying the making of values*. In Dussage, I., Helgesson, C.-F. and Lee, F. (eds) *Value Practices in the Life Sciences and Medicine*. Oxford: Oxford University Press.
- Dussage, I., Helgesson, C.-F., Lee, F. and Woolgar, S. (2015b) On the omnipresence, diversity, and elusiveness of values in the life sciences and medicine. In Dussage, I., Helgesson, C.-F. and Lee, F. (eds) *Value Practices in the Life Sciences and Medicine*. Oxford: Oxford University Press.
- Hall, S.S. (2006) *Size Matters: How Height Affects the Health, Happiness, and Success of Boys – and the Men They Become*. New York: Houghton Mifflin Company.
- Jutel, A. (2009) Sociology of diagnosis: a preliminary review, *Sociology of Health & Illness*, 31, 278–99.
- Jutel, A. and Nettleton, S. (2011) Towards a sociology of diagnosis: reflections and opportunities, *Social Science & Medicine*, 73, 793–800.
- Lamont, M. (2012) Towards a comparative sociology of valuation and evaluation, *Annual Review of Sociology*, 38, 201–21.
- Latimer, J., Featherstone, K., Atkinson, P., Clarke, A., et al. (2006) Rebirthing the clinic: the interaction of clinical judgement and genetic technology in the production of medical science, *Science, Technology & Human Values*, 31, 599–630.
- Lee, N. (1999) The challenge of childhood: distributions of childhood's ambiguity in adult institutions, *Childhood*, 6, 455–74.

Lowy, I. (2015) Norms, values and constraints: the case of prenatal diagnosis. In Dussage, I., Helgesson, C.-F. and Lee, F. (eds) *Value Practices in the Life Sciences and Medicine*. Oxford: Oxford University Press.

Mol, A. and Elsmann, B. (1996) Detecting disease and designing treatment. Duplex and the diagnosis of disease leg vessels, *Sociology of Health & Illness*, 18, 609–31.

Morrison, M. (2008) *Beyond therapy? Investigating biomedical enhancement in the case of human growth hormone*. Unpublished Doctorate. Nottingham: University of Nottingham. 11

Morrison, M. (2015) Growth hormone, enhancement and the pharmaceuticalisation of short stature, *Social Science and Medicine*, 131, 305–12. 12

Muniesa, F. (2012) A flank movement in the understanding of valuation. In Adkins, L. and Lury, C. (eds) *Measure and Value: Sociological Review*. Oxford: Wiley-Blackwell.

Nordlund, C. (2015) The moral economy of a miracle drug: on exchange relationships between medical science and the pharmaceutical industry in the 1940s. In Dussage, I., Helgesson, C.-F. and Lee, F. (eds) *Value Practices in the Life Sciences and Medicine*. Oxford: Oxford University Press.

Roscoe, P. (2015) A moral economy of transplantation: competing regimes of value in the allocation of transplant organs. In Dussage, I., Helgesson, C.-F. and Lee, F. (eds) *Value Practices in the Life Sciences and Medicine*. Oxford: Oxford University Press.

Sandberg, D.E. and Colman, M. (2005) Growth hormone treatment of short stature: status of the quality of life rationale, *Hormone Research*, 63, 275–83.

Tanner, J.M. (1981) *A History of the Study of Human Growth*. Cambridge: Cambridge University Press.

Valuing height: Diagnosis, valuation, and the case of idiopathic short stature: Supplementary material

Guide to Interview Respondents

In the interests of clarity, direct quotes from interviewees are identified by a short code of the form 'UK X' or 'NAM X'. The first part of the code identifies the interviewee as working in either the UK or North America and the 'X' represents an identifying number given to each interview transcript.

Descriptive details of each respondent are provided below along with the appropriate code. Note that in the original study endocrinologists who worked only with adult patients were also interviewed. These respondents did not discuss idiopathic short stature and are not cited in the main manuscript or listed in this table.

CODE	RESPONDENT DESCRIPTION
UK 1	UK-based Professor of Paediatric Endocrinology with experience of working with pituitary and recombinant growth hormone formerly involved with the Dept of Health GH Committee.
UK 3	UK-based Reader in Paediatric Endocrinology involved with research on GH Therapy and physiology. Some experience of pituitary GH, mainly worked in recombinant era.
UK 4	UK-based Professor of Paediatric Endocrinology with extensive experience of pituitary as well as biosynthetic GH and the UK Medical Research Council / Dept. of Health committees.
UK 5	UK-based Clinical Senior Lecturer in Child Health with experience of current therapeutic practice using recombinant GH.

NAM 2	US-based Professor of Paediatrics with experience of pituitary and recombinant GH.
NAM 3	US-based Professor of Paediatrics with extensive experience of pituitary GH and the National Pituitary Association as well as recombinant GH. Also involved in early investigations of GH in ageing.
NAM 4	US-based Professor of Paediatrics, Pharmacology and Therapeutics, with clinical experience of growth hormone in children and adults and research interests on GH use in ageing.
NAM 5	US-based senior figure with university medical school affiliated charitable foundation for children's health. Some experience of pituitary GH, mainly worked in recombinant era in both treatment and extensive research on molecular basis of growth pathology.
NAM 6	US-based Associate Professor in Child Behavioural Health with extensive experience in psychosocial aspects of short stature.
NAM 8	Canada-based Professor and Chair of Paediatrics, experience of Canadian pituitary hormone era and recombinant hormone.

Table 1.0: Details of interview respondents discussing idiopathic short stature.