Chapter 2
Uncertainty and Clinical Method

Trisha Greenhalgh

Prologue

A Friday evening general practice surgery. Four cases, part-fictionalised to protect confidentiality.

Case 1. Lindsay, a 36-year-old woman with a breast lump. She is 7 weeks pregnant with her second child. She had a benign lump removed from this part of the same breast 4 years ago and was told to get any further lumps checked out promptly. She stopped breast feeding 2 months ago. I examine her. She has tender breasts and an ill-defined lump in the upper outer quadrant of the left one, but no ‘red flag’ signs (such as skin tethering or enlarged lymph nodes).

Case 2. Chris, a 54-year-old man with an itchy rash. For 3 days he has had about ten reddish, blistering and encrusted spots on the left side of his chest and back. He also has three or four on his left thigh and some on both sides of his forehead. If the rash were confined to his trunk, it might be shingles, since it curves down in a crescent shape, but given the wider distribution, it’s more likely to be insect bites. He apologises for being here at all. His wife thought it might be an illness, so she made an appointment for him. I am surprised to smell alcohol on Chris’s breath, even though it is only 4.30 p.m. A pop-up prompt from my computer reminds me that Chris is a current smoker and that his Framingham risk score for developing cardiovascular disease in the next 10 years is 15%.
Case 3. Aisha, a 63-year-old British Afghan woman with a painful hip. She has come to tell me about her experiences in the orthopaedic department. She has had X-rays, blood tests and scans. A firm diagnosis of advanced osteoarthritis has apparently been made. Aisha has been advised to have a total hip replacement as soon as possible. She asks me what I think. She is an educated woman and fluent in English, but she politely rejects my offer of a leaflet and video giving the facts about total hip replacement. It’s not the details of the operation she is interested in but its relevance to her own situation. I ask what issues are uppermost in her mind. She explains how much she hates hospitals, how her case was dealt with solely by a doctor in training rather than the renowned consultant hip surgeon to whom I had referred her and how her children (who work in medically related professions) feel that she should have the operation. She has been on the Internet and found out that the small dose of diclofenac which she takes at night is ‘bad for you’, but the other painkillers we have tried are less effective.

Case 4. Faisal, a 24-year-old Iranian man who has booked a double appointment and attends with a professional interpreter. He thinks he has wax in his ears. I check them, confirm the wax and explain through the interpreter that he must put in drops and then see the nurse for syringing. As the interpreter gets up to leave, he hesitates and mumbles something. The interpreter explains that Faisal says he is depressed. On questioning, he has been depressed for 5 years. He doesn’t want to talk about the reasons. A letter from the Medical Foundation for Care of Victims of Torture was scanned onto his electronic record 3 years ago, but I cannot open it because of a technical fault. Faisal looks at the floor and makes no attempt to elaborate on his story despite sharp words from the interpreter (‘I’ve told him to answer you, doctor’). I reflect on whether Faisal would be more forthcoming with a different interpreter or a different doctor. The interpreter appears uncomfortable with my apparent indecision and suggests I put him on Prozac.

I had gone to surgery last Friday intending to collect some cases to illustrate uncertainty in clinical practice, having planned to start writing this chapter over the weekend. When I returned, I told my husband I was disappointed: not much uncertainty had been evident today. I would have to try again next week. Yet when I wrote out each of my 15 encounters as a brief case history, uncertainty was a central feature of every single one. I wondered why I had been blind to this at the time, even though I had approached the day’s clinic specifically looking for examples of uncertainty.

Introduction: Different Perspectives on Uncertainty in Clinical Practice

General practice sits between what Mishler, following Habermas, called the voice of medicine (i.e. the assumptions, values and logic of a clinico-pathological world view) and the voice of the lifeworld (i.e. the assumptions, values and logic of a person-living-in-society world view) (Habermas 1981; Mishler 1984). The general
practitioner, and any other clinician who manages illness in a community context, must be highly trained – up to date with the relevant literature and technically competent – and use this knowledge and skill set in clinical decision-making and care delivery. He or she must also possess the personal qualities that support strong therapeutic relationships – especially the ability to connect at a personal level with the patient and engage with their story.

Uncertainty in clinical method is inherent both to the voice of medicine (the science of clinical decision-making in relation to diseases) and to the voice of the lifeworld (the elements of clinical practice to do with the management of illness). Both diseases (those textbook entities which we seek to diagnose, treat, prevent and monitor) and illnesses (the lived experiences and symbolic meanings of sickness, healing, coping and dying) unfold in ways we can neither fully predict nor fully control.

Thirty years after Habermas described his ‘system’ and ‘lifeworld’ distinction, it is time to add a third voice: the voice of technology – the opportunities, constraints and organisational logic engendered by the ubiquitous use of computers and other information and communication technologies. Paperless consulting, touch-screen self-registration, electronic templates for chronic disease monitoring, Internet-based guidelines and protocols, computer-generated risk scores, online appointment booking, automated data transfer (records, laboratory tests and discharge summaries), telephone consultations, text message reminders, downloadable patient leaflets and digital financial flows are all part of the business as usual of general practice – as is the periodic failure of all these technologies and the idiosyncratic ways in which humans use them. Both staff and patients in the contemporary health care organisation need to be digitally astute and skilled in working around these limitations.

Ironically, given the promise of technology to provide accurate, reliable data at the touch of a button, uncertainty is a particular feature of multi-professional care that is supported by distributed information systems (e.g. a shared computer network which allows different staff members to call up and work on a patient record from their terminal, with automated links to pathology, radiography and outpatient booking services). Someone on a multi-professional care team may know an item of information, and/or a data item might be filed somewhere on the system, but because of the sheer size and complexity of the socio-technical system, that item may not be accessible or locally meaningful to the member of the team who is currently seeing the patient. As we shall see, the voice of technology does not correspond unproblematically to the harmonised voice of the multi-professional team.

In this chapter, I will review the literature on clinical uncertainty, touching on philosophical, theoretical and empirical perspectives, but I will focus mainly on my own clinical perspective – that is, the perspective of a doctor who sees patients, reflects on my practice and strives to make my contribution to a wider team effort. In this chapter, I will summarise four approaches to uncertainty that I find

---

1 I am indebted to my general practitioner Ph.D. student, Deborah Swinglehurst, for many discussions on ‘voice’ in the electronic patient record.
particularly relevant to front-line clinical work and especially to clinical judgement. These perspectives overlap considerably but, at a philosophical level, are not fully commensurable with one another.

They are (a) evidence-based medicine, which considers uncertainty in probabilistic (Bayesian) terms; (b) narrative medicine, which considers uncertainty in terms of the open-endedness of the story form and the creative space in which storytelling is performed; (c) medicine as case-based reasoning, which considers uncertainty in terms of situated ethical judgements and tacit knowledge (and which, I feel, overcomes many of the limitations of evidence-based and narrative-based medicine); and (d) multi-professional care, which considers uncertainty in terms of how knowledge, reasoning and action are spread across a network of people and technologies.

I acknowledge that these perspectives do not represent a definitive or complete taxonomy (the psychodynamic approach advocated by Michael Balint might be considered a separate category, though some would classify it as an application of narrative medicine; and a ‘critical theory’ approach might consider uncertainty from a political power perspective). The four I have chosen might serve as ideal types against which other approaches might be compared and contrasted.

In this chapter, I have used the term ‘clinical method’ synonymously with ‘clinical judgement’. I did this unconsciously, perhaps reflecting the approach that prevailed in the mid-1980s when I first studied general practice and began to be inspired by writers like George Engel, Marshall Marinker, Iain McWhinney and Eliot Mishler (1984; 1978; 1983; 1986). Nowadays, clinical method has taken on a somewhat different meaning, referring not to the situated judgements made by a clinician about a real patient in a real predicament but to a set of transferable clinical skills and task-based procedures (such as how to examine an abdomen), especially when taught to students. When ‘clinical method’ is given this more contemporary meaning, it is possible to assess it in objective structured clinical examinations (OSCEs) and even do randomised trials in which it is the dependent variable (Murray et al. 1997; Al-Dabbagh and Al-Taee 2005).

When this different meaning of the term was pointed out to me by the editors, I contemplated changing ‘clinical method’ to ‘clinical judgement’ throughout this chapter, but after some reflection, I chose to keep the terminology as I had originally used it. To conflate clinical method with context-free performance in an OSCE is an act of reductionism from which academic general practice urgently needs to be redeemed. I want to reclaim the original meaning of the term, and I hope this chapter will illustrate why.

Clinical Method as Evidence-Based Decision-Making

Evidence-based medicine (EBM) is ‘the use of mathematical estimates of the risk of benefit and harm, derived from high-quality research on population samples, to inform clinical decision-making in the diagnosis, investigation or management of individual patients.’ (Greenhalgh 2010). The expression ‘evidence-based medicine’
is a relatively new packaging of epidemiological concepts that have been around for decades. EBM was introduced by Sackett and colleagues as a reaction to what they called ‘decision-making by anecdote’ and which might be more neutrally depicted as replicating the iterative hypothesis-testing of expert diagnosticians (Kassirer et al. 1982).

Proponents of EBM have pointed out the folly of assuming (as we did for centuries) that generations of doctors would become competent merely by watching other doctors make decisions and argued that a more systematic approach to incorporating best research evidence into clinical judgement would reduce inappropriate, wasteful and sometimes overtly harmful variations in practice (Eddy 2005).

Epidemiologists, who generate the science on which EBM is based, look at diseases in populations. They begin with a sample – a systematically collected group of people who either have a disease or who might develop it – and they follow this sample to see what happens to them. Sometimes, the sample is allocated at random to receiving treatment A or treatment B, and the outcomes (benefits and harms) compared between the two groups: the much-feted randomised controlled trial. Clinical epidemiology thus generates numbers that inform decisions in prevention, diagnosis, prognosis, treatment and surveillance – such as the sensitivity and specificity of a diagnostic test or the ‘number needed to treat’ (NNT), an estimate of the efficacy of a drug. This approach rests firmly on objectivist philosophical principles (a) that there is a physical reality independent of the observer; (b) that there are things to be measured and facts to collect; and (c) that the purpose of research is to find valid, reliable, transferable truths (‘empirical facts’) that can be used to inform decisions by other people in other settings.

From a philosophical perspective, EBM is sometimes considered to be the last bastion of crude empiricism, though Mike Kelly and Tessa Moore have recently argued, based on the philosophy of Hume and Kant, that EBM rests on sophisticated (albeit largely unexamined) reasoning which precedes empirical observation (Kelly and Moore 2012). In their words:

> The principles of the elimination of the possibility of bias in the hierarchy of evidence, of the rule-driven principles of guideline development and appraisal are based on an ideal version of the scientific method, which owe more to the logical precepts of the a priori relations of ideas than they do to messy empirical observation. (p. 10)

When I confirmed the lump that Lindsay had found in her breast, my chain of reasoning included asking myself ‘if I took a sample of a thousand women in their 30s, each of whom had a breast lump, in what proportion of them would the lump turn out to be cancerous?’ (an estimate of prevalence). Importantly, the answer to this question depends on where the sample of women came from. If the sample were drawn from a specialist breast clinic (say, 1,000 consecutive referrals of breast lumps from general practice), around a quarter of those women would have cancer. But if I took a thousand women in their 30s attending general practice with a breast lump, only around one in 40 would have cancer (Greenhalgh 2010).

Epidemiological research has greatly improved our ability to predict whether a particular patient is likely to develop a particular disease. By following thousands of women prospectively for many years, and carefully recording key data items over
time, they have produced sophisticated risk scores in which different factors (a history of a benign breast lump, having a first child late in life, past use of hormonal contraceptives and so on) contribute different amounts to a probability estimate for the development of cancer. This kind of research helps us explain, for example, why Japanese women are less likely to develop breast cancer than British women. It has also informed an emerging science of clinical prediction rules and scoring systems which offer the best combination of symptoms, signs and laboratory tests to ‘rule in’ or ‘rule out’ a particular diagnosis or risk state and which, once validated, can be used to help resolve a host of clinical conundrums from diagnosing the cause of a red eye to estimating the risk of falls in an older person (Reilly and Evans 2006).

A clinical prediction score will not tell me whether Lindsay has cancer, but it will reduce the uncertainty around this question. By combining everything I know about her risk factors and the results of my clinical examination, I can estimate that her chance of having breast cancer is rather more than 2.5% (the average level for someone her age seen in general practice). She needs to be referred to a breast clinic, where she is likely to be offered a combination of three tests – examination by a specialist (whose skill in distinguishing benign from suspicious breast lumps on clinical grounds is likely to be greater than mine), an ultrasound examination and a fine needle biopsy (a fourth test, a mammogram X-ray, will probably be omitted because of her pregnancy). Each of these tests alone is incapable of ruling breast cancer in or out with 100% accuracy, but the combination of all three ensures that once all the results are available, Lindsay’s doctors will be able to say with confidence either that the chance of her having cancer is vanishingly small or that she almost certainly does have cancer or that there is sufficient residual doubt to justify a prompt operation to remove the lump so that a firm histological diagnosis can be made.

Many readers, especially those who have been trained as doctors since the 1990s, will already be familiar with the principles of EBM and how probabilities derived from epidemiological research are used to produce guidelines and protocols (and, sometimes, programmed into computerised decision support tools) so as to support a scientific, objective and rational approach to clinical decision-making (Eddy 2005). In the language of EBM, a diagnostic test should be sensitive (i.e. it picks up all or very nearly all people who have the disease) and specific (i.e. it excludes all or nearly all people who do not have the disease). A treatment (e.g. drug, operation or lifestyle change) should produce a high level of benefit (e.g. cure or extend life) but a low level of harm (i.e. no serious side effects).

Some diagnostic tests perform extremely well at reducing uncertainty. Take the CAGE questionnaire for example. Faced with a patient who has a possible alcohol-related problem, ask four questions: (a) have you ever felt you needed to Cut down on your drinking? (b) have people Annoyed you by criticising your drinking? (c) have you ever felt Guilty about drinking? and (d) have you ever felt you needed a drink first thing in the morning (Eye-opener) to steady your nerves or to get rid of a hangover? In one study, a CAGE score above or equal to 2 had a sensitivity of 93% and a specificity of 76% for diagnosing a ‘problem drinker’ (Berandt et al. 1982).
I thought of using this simple score with Chris, but decided against it. I will explain this decision from a narrative medicine perspective in the next section, but to justify it from an EBM perspective, reassuring evidence is emerging that strictly applying population-based evidence to individual cases produces fewer overall benefits (and more overall harms) than a more nuanced and flexible approach, even if the latter appears less ‘evidence-based’ (Kent and Shah 2012). Similarly, some treatments are extremely effective and have few downsides. An example is triple therapy (two antibiotics and a proton pump inhibitor) for peptic ulcer disease. For every 1.1 patients given triple therapy, the causative bacterium (helicobacter) will be eradicated in one, and for every 1.8 patients treated, one will remain cured of peptic ulcer 1 year later. Whilst minor side effects with triple therapy are common, serious ones are extremely rare. Thus, whereas 20 years ago there was much uncertainty about how best to treat peptic ulcer, there is now strong consensus (Moore 2009).

However, many diagnostic tests are neither sensitive nor specific, and very few treatments are both highly beneficial and lacking in serious side effects. Herein lies a major – and arguably ineradicable – source of clinical uncertainty. In relation to Faisal, for example, even putting aside whether a diagnosis of depression is either justified or helpful (see section “Clinical Method as Narrative”), epidemiological research has shown that whilst selective serotonin reuptake inhibitors (SSRIs, the class of drug to which Prozac [fluoxetine] belongs) may lead to significant reduction in the symptoms of depression, they also often lead to serious side effects, including an increase in suicidal tendencies and long-term dependence (Cipriani et al. 2010). Far from informing a clear course of action, the epidemiological evidence base on SSRIs merely serves to confirm that the advantages of putting Faisal on this drug may or may not be outweighed by the disadvantages.

The EBM community has recognised that in many if not most clinical decisions, there are pros and cons of different options and that different patients may place different values on each option. Much work has been done to develop ways of representing epidemiological research findings diagrammatically so as to support informed, shared decision-making by patients (Han et al. 2011a; Edwards and Elwyn 2010). Aisha, for example, may prefer to tolerate moderate pain and disability than succumb to the surgeon’s knife; others with the same symptoms would be keener for an operation.

Would Aisha’s uncertainty be helped by formal ways of mapping benefits and risks and asking her to rate her preferences on a scale of nought to five? Perhaps, though I may have to work on her to increase her receptivity to such a task. Working with patients to share uncertainty and address it systematically is not something many clinicians do routinely, though the abstract books of academic conferences are replete with reports of ‘tools’ intended for use in the clinical consultation for applying essentially Bayesian (probabilistic) reasoning to some clinical scenario or other. The difference between the enthusiasm with which such tools are produced and the reluctance of clinicians to use them in practice would surely be a worthy topic for further research. Indeed, a Ph.D. student of mine had a go at this with diabetes risk scores (Noble et al. 2011).
Let us return to Faisal’s unhappiness. A rapidly expanding branch of epidemiology known as ‘personalised medicine’ seeks to square the circle between population-derived evidence (which can tell us, at best, whether fluoxetine reduces depression symptoms on average in people like Faisal) and the personal benefit-risk equation in a particular patient (will fluoxetine reduce depression symptoms in Faisal, and also, of the many antidepressant drugs available, which is best suited to his particular metabolism?). The epidemiology of personalised medicine (not to be confused with ‘evidence-informed individualised care’ – that is, the judicious combination of scientific reasoning with deep and intimate knowledge of one’s patient (Miles and Loughlin 2011)) works by collecting genetic information and tissue samples as well as conventional risk factors over time, thereby (it is hoped) narrowing the confidence intervals around a prediction of risk or benefit for an individual. The UK Biobank project, for example, has already recruited half a million people and aims to follow them for 20 or more years, leading eventually to several thousand data items for each person (Hewitt 2011).

With the help of such large databases and powerful computer programmes, it is theoretically possible (so the protagonists of personalised medicine believe) that we will reach a point where instead of ‘on average’ estimates of the likelihood of developing certain conditions or of the benefit or harm of different treatment options, we will be able to say to patients like Lindsay ‘your breast cancer is of xx genetic type; given your personal genetic and metabolic profile, the chance of achieving cure with the best treatments currently available is 89%, but the chance that your cancer will subsequently recur is 63%’. Such apparently accurate predictions (a future possibility, some say) have the potential to reduce uncertainty – but they are unnerving and raise important ethical and legal questions about privacy, ownership of data and the regulation of testing (Bourret et al. 2011).

It is beyond the scope of this chapter to pursue the finer points of clinical epidemiology, evidence-based medicine, shared decision-making or the promise of personalised medicine. But it should be clear from this brief account that this approach to clinical method rests on a number of assumptions (Henry 2010; Braude 2009; Djubegovic et al. 2009; Mol 2008). First, it assumes that diseases and risk categories exist and are useful entities – that is, that it is both possible and helpful to sort patients into those who do or do not fulfil the diagnostic criteria for a disease or pre-disease state (breast cancer, peptic ulcer, high cardiovascular risk, ‘problem drinking’ and so on).

Second, it assumes that clinical method is wholly or mostly to do with the epidemiology-informed tasks of preventing, diagnosing, treating and monitoring disease, and hence with making clinical decisions (should I do this test, that test or no test? Should I prescribe this drug, that drug or no drug? Should I advise my patient to make this or that change to their lifestyle? and so on). Third, it assumes that evidence collected in some other population sample at some other time is more or less transferable to the patient in front of you. Fourth, it assumes that this information (typically derived from a randomised controlled trial or longitudinal cohort study) is at least as valuable as, and should sometimes override, the tacit, intuitive knowledge of the experienced clinician (see section “Clinical Method as Narrative”).
Fifth, it assumes that facts and values are separable and that patients’ values and preferences are quantifiable, stable over time and readily linked to the diagnostic or treatment options in the clinical decision tree. Finally, the expansion of the evidence base to include genetic material and an increasingly detailed environmental risk profile on hundreds of thousands of research participants rests on the (arguably highly questionable) assumption that this complex system will exhibit linear dynamics, making accurate prediction of individual risk and benefit possible.

These assumptions hold true – or at least, they approximate so closely to the truth that for all practical purposes we can ignore the difference – in some but not all situations. The uncertainty in Lindsay’s case (Does she have breast cancer? Should her lump be removed urgently?) and perhaps Aisha’s case (What are the benefits and risks of medical and surgical management of moderate osteoarthritis of the hip?) may be substantially reduced by examining the problem through an EBM lens. But the uncertainty in relation to Chris’s skin rash will not, since the evidence base for mild, non-specific skin rashes is all but non-existent (Perhaps because virtually all such rashes are self-limiting and this problem would be a low research priority).

And in Faisal’s case, where the uncertainty is particularly pressing (What happened to him? Does he have depression? Is depression a helpful or unhelpful category to describe his illness? What are the available non-drug options? Who else is around in his life? Why is his relationship with this professional interpreter so awkward? and so on), only a fraction of this uncertainty will be reduced by recourse to the epidemiological research literature. Furthermore, even assuming a future state in which we will know a lot more than we currently do about the genetics of SSRI response, only the most confident pharmaco-epidemiologist would argue that a full genetic profile will provide the key to eradicating the uncertainty in Faisal’s case.

A recent review article in the journal Medical Decision-Making offered a ‘new conceptual taxonomy’, in which the authors recognised three sources of uncertainty in health care: (a) probability (indeterminacy of future outcome), (b) ambiguity (lack of evidence, contested evidence or imprecision in estimates), and (c) complexity (multiplicity of causal factors and interpretive cues) (Han et al. 2011b). This seems a reasonable way to parse the types of uncertainty which fall within the evidence-based medicine paradigm, but it does not engage to any significant degree with the three other perspectives on uncertainty covered in this chapter.

Clinical Method as Narrative

If evidence-based decision-making is the objectivist, disease-oriented perspective on clinical method, narrative medicine is usually thought of as the subjectivist, illness-oriented perspective. Whether a disease or pre-disease state ‘really’ exists as an objective entity or not, when we are ill or at risk (or believe ourselves to be so), we must make sense of our experience and learn to cope with it via what Mike Bury has called ‘rebuilding spoiled identity’ (Bury 1991). The illness journey and efforts to manage it unfold against a rich tapestry of symbols, actions and events in the
wider world; efforts to manage illness (by patient and professional) have social, cultural and moral dimensions that vary in different contexts and at different times. Narrative medicine, whose cross-disciplinary roots include literary studies, sociology, phenomenology and psychoanalysis, is the art of engagement with the patient as narrator of their personal story and the use of this ‘storywork’ as a therapeutic force (Charon and Montello 2002; Launer 2002).

Chris has attended my clinic smelling of alcohol. Is he a ‘problem drinker’? Whilst one approach to reducing the uncertainty in Chris’s case might be to invite him to complete the CAGE questionnaire described briefly in the previous section, another might be to be open to Chris’s story in his own words. He tells me that he has had to take a half-day’s annual leave to keep this appointment (which had been made by his wife on his behalf) and that rather than ‘waste’ this enforced leave, he has been watching the football on TV. Intuitively, I compare this story fragment with my stock of past experience of patients who have attended smelling of alcohol (for innocent or pathological reasons) and with my stock of experiences of family and friends who watch the football when on annual leave. Rightly or wrongly, I decide that I have a plausible enough narrative to make CAGE redundant. Perhaps the issue here is not whether prescribed ‘good practice’ (in this case, the use of an evidence-based assessment instrument) is used but my judgement as an experienced practitioner about whether this instrument is a ‘good fit’ with the emerging dialogue between patient and clinician. That judgement – the essence of what is often referred to as ‘narrative [based] medicine’ – is aesthetic, ethical and (above all) intuitive.

The narrative approach originated within, and remains most widely used and researched, in relation to mental health is the so-called talking therapies. Especially in relation to patients whose unhappiness or confusion is linked to the contexts and complexities of their lives, many authors have questioned whether it is useful to manage illness by assigning a disease category (‘depression’, ‘anxiety’, ‘post-traumatic stress disorder’) and linking treatment to those categories (Launer 2002; Heath 1998). Rather, these authors propose, the clinician’s central role may be to serve as the audience for the story – that is, as the active listener whose concern and questioning shape the narrative and work towards it acquiring coherence and purpose.

The skilled narrative clinician notes the genre of the story (perhaps using Art Frank’s widely cited taxonomy of illness narrative – restitution, tragedy, quest, chaos (Frank 1995)), its characterisation (does the patient depict him or herself as hero, victim, bystander etc.; is the surgeon a hero or villain; do family members play supportive or destructive roles?), its setting (is the story primarily set in the home, the workplace, the community etc.?), the key actions and events (what happens – and by whose agency?) and how a plot is woven through the use of literary devices (metaphor, imagery, tropes). Thus, the clinician learns how and in what way the person is ill and, by active listening, helps co-construct a healing or coping narrative (Greenhalgh and Hurwitz 1999).

Physical illness (including chronic illness and cancer), while often readily classifiable as disease, can also be thought of as primarily a challenge of sense-making and identity [re]construction (Bury 1991; Frank 1995; Mattingly 1998). Frank has
written powerfully about his own cancer journey and the crucial role of the clinician as witness to the suffering of the individual with ‘deep illness’ (Frank 1998). My own team recently showed that the narratives of people with diabetes can be analysed in two ways, revealing both a biomedical perspective (the various tasks of self-management including disease monitoring, planning and monitoring the diabetes diet, managing medication, taking exercise, accessing health care, looking after feet and so on) and also a narrative perspective in which these ‘biomedical’ tasks acquire social meaning and moral worth in the patient’s ‘lifeworld’ (Greenhalgh et al. 2011). ‘Self-management’ is as much about socio-emotional work as it is about monitoring biomarkers (Hinder and Greenhalgh 2012).

Engagement of the health professional with the patient’s narrative is thus not an alternative to considering the biomedical aspects of care, but a means by which the science of caring for the person who is ill comes to make sense to both patient and clinician. This engagement also has an ethical dimension, representing as it does the professional and human commitment to acting in the patient’s interests. As Charon and Montello have put it, ‘the singular case emerges only in the act of narrating it and duties are incurred in the act of hearing it’ (Charon and Montello 2002).

Aisha’s case shows not only how a narrative approach is relevant to the entirety of the illness journey (the unfolding of the over-arching story of worsening arthritis – a story which will be built and understood through continuity of care), but also how it may help reduce uncertainty at particular steps in that journey. In this particular consultation, Aisha faces a decision: to undergo the hip replacement operation or not. But she is clear that she does not want to know any more facts and figures about the operation, which is why I am not convinced that a ‘shared decision-making’ framework will work in her case. Perhaps, we should explore how the pain and disability of an arthritic hip are affecting her role in her family and community. Perhaps, I should invite the story of how Aisha came to hate hospitals or of how her grown-up children have responded to the unfolding of her illness journey. Thus, through dialogue, Aisha and I might co-construct a narrative of today’s plight in relation to her chronic illness.

Faisal’s case highlights how a purely ‘subjectivist’ narrative medicine approach isn’t going to get us very far with complex cases. For sure, Faisal (newly registered with my practice) has a story to tell, but he doesn’t want to tell it. Faisal’s asylum seeker status, the letter from a charity specialising in victims of torture, his averted eyes and burdened body language all suggest the need for intimacy, trust, continuity of care, a common vocabulary and an ethical commitment on the part of the listener as preconditions for beginning his storywork. Yet the interpersonal dynamics of this awkward tripartite consultation provide none of these contextual conditions. My role as clinician is not to attempt to extract his painful story with the assistance of this interpreter (who, for all her professional qualifications and shared cultural background, appears oddly disengaged from Faisal’s predicament) but somehow to bear witness to his suffering despite her presence. I must also work within the constraints of the here and now to create some possibilities for a more productive consultation sometime in the future.
It is not Faisal’s untold story (narrative as noun) that is the issue here but the dynamics of the telling – and the not telling (narrative as verb). The Russian philosopher and literary critic Michael Bakhtin emphasised that whenever a story is told, it is shaped by the listener or (imagined) reader, since every utterance is made in response to, or in anticipation of, the utterance of another human being (Bakhtin 1981). Contemporary narrative scholars have moved beyond a narrow focus on the text (i.e. on studying the structural elements of the story) to consider the teller-listener relationship and the act of storytelling (Riessman 2008). In this performative framing, narrative medicine is not a purely subjectivist phenomenon. Rather, the story is a practical accomplishment, a product of the dynamic, teller-listener interaction. And it is in this performative space rather than in the narrative itself that the scope for resolving uncertainty lies (Charon and Montello 2002; Mattingly 1998; Malterud 2006).

A number of related streams of academic work are worth mentioning here. First, there is work on the therapeutic relationship, including Frankel’s notion of ‘relationship-centred care’ (Frankel 2004) and earlier seminal work on the same theme by Michael Balint (1957) and Carl Rogers (1951). In contrast to approaches which emphasise tasks and processes, these authors have argued that relationship conditions offered by the clinician (empathy, congruence and unconditional positive regard) are in and of themselves therapeutic and that this relationship creates the preconditions for effective communication (perhaps, co-construction of a narrative). Balint encouraged general practitioners to reflect in groups on clinical cases, sharing stories of the stories they had co-constructed with their patients and using their fellow practitioners as a critical audience to suggest potential therapeutic options.

**Clinical Method as Case-Based Reasoning**

It is worth reflecting here on why, as my Friday evening clinic unfolded in real time, I experienced little in the way of conscious feelings of uncertainty. This is at least partly because as an experienced clinician working in a familiar setting and most of my patients were known to me, my reasoning was predominantly intuitive (Greenhalgh 2002). That is not to say I was especially knowledgeable about the clinical topics I was encountering, but that I was processing multiple sources of information rapidly and largely unconsciously – and doing what normally works (Dreyfus and Dreyfus 1986; Benner 1984).

As Kathryn Montgomery has shrewdly observed, ‘Clinicians’ lack of curiosity about clinical thinking (with its concomitant appeal to science) turns out to be characteristic of all practice, an apparently unavoidable consequence of the requirement that practitioners act despite uncertainty’ (p. 193) (Montgomery 2009). As clinical practice becomes ever more rationalised and protocol-driven, and as professional development shifts further from reflecting on the curious features of individual cases in favour of achievement of specific learning targets, the question arises: are we in danger of retreating even further from the intuitive side of our expertise?
In relation to Lindsay’s case, for example, it was only afterwards that I became conscious of a memory of a lecture on breast cancer in pregnancy some 30 years ago at medical school. In it, we had been reminded that whilst breast cancer in pregnancy is rare, when it occurs it is rapidly progressive, and we should therefore refer any suspected cases urgently (Amant et al. 2012). This snippet of wisdom from an experienced (and probably now long dead) professor of surgery was overlaid by more recent and more conscious memories of evidence-based guidelines for referral of breast lumps in non-pregnant women – and by the rhetorical power of the story within a story of how Lindsay had previously been advised to present without delay should a lump recur in this part of her breast.

The notion of tacit knowledge is relevant here. As Polanyi observed, tacit knowledge is embodied, tied to individuals, impossible to codify or measure objectively and hard to transmit to others (Polanyi 1962). It functions at the periphery of attention and – in Henry’s words – ‘forms a largely taken-for-granted foundation that makes the information on which the clinician focuses directly, such as the patient’s story and the significance of her [clinical examination], possible’ (Henry 2010). This is the knowledge that comes from years of walking the wards, consulting in clinic and visiting patients in their homes – and above all, the knowledge which is built by accumulating cases (Cox 2001). It is this, rather than an encyclopaedic knowledge of all the relevant guidelines, which defines the difference between the expert and the novice clinician. And it is this tacit dimension, necessarily accessed intuitively, which informs that elusive judgement on what is the right thing to do for Lindsay, Chris, Aisha and Faisal.

That the knowledge needed to enact the clinical encounter is often tacit, context-bound and ephemeral rather than codifiable, transferable and enduring often goes under-recognised and under-explored by academic commentators on clinical method. In Faisal’s case, for example, I made a judgement not to pursue his symptoms of depression exhaustively in this particular encounter. Rather, I chose to try to convey to him that I was open to developing a therapeutic relationship and working on his narrative at some other time in the future (perhaps with a different interpreter). Designers of electronic records may miss the crucial importance of such situated communication and produce artefacts (such as pop-up prompts or inflexible templates in electronic records) that fit poorly with the subtle micro-detail of clinical work.

The ubiquitous pop-up prompt, whilst driven by ‘evidence-based’ guidelines, complicates and potentially disrupts the delicate interpersonal interaction on which the clinician-patient consultation has traditionally depended (Balint and Norrell 1983). It is difficult enough to build a therapeutic relationship with Faisal across a language barrier, especially given the limitations of his illness, my own limited understanding of his circumstances and the nature of the cultural support on offer, but this would surely be made even more difficult if Faisal knew I was being prompted (and financially incentivised) to ‘offer Chlamydia testing’. Moreover, I deliberately did not enter Faisal’s ‘depression’ as a new coded item on his electronic record. Had I done so, a template would have been triggered which would have included a structured questionnaire on mental health symptoms. Failure to get Faisal
to complete that questionnaire and enter his score in the appropriate box would have led to my practice incurring a financial penalty in the pay-for-performance Quality and Outcomes Framework used in the UK (Roland 2004).

Whilst inscribing evidence-based guidelines in technology is not in and of itself a ‘bad thing’, the voice of technology can be a two-edged sword when subtle emotional work is being done in the consultation. The technological voice must be actively and reflexively managed (and where appropriate, subverted or worked around) as part of case-based reasoning rather than allowed to direct the encounter whatever else is going on in it (Swinglehurst et al. 2011).

Evidence-based and narrative-based approaches to clinical method may on one level be analytically separated (and based on very different philosophical assumptions), but they are not mutually exclusive nor should they be considered in a simplistic zero-sum relationship in which more of one presumes less of the other. There is no evidence that clinicians who try to practise in an evidence-based way necessarily undergo attrition of their humanitarian values nor that the use of evidence-based guidelines and protocols produces a ‘science-based meritocracy on the patient ward’ (Timmermans and Angell 2001). In Aisha’s case, a skilled application of clinical method would combine the narrative approach to reducing uncertainty (e.g. drawing out her story of why she is so opposed to hip surgery) with an evidence-based approach to optimising medical management of her pain and disability.

Those who seek a philosophically commensurate alignment of evidence-based and narrative-based clinical methods should read Montgomery’s excellent book ‘How Doctors Think’ (Montgomery 2006). Drawing on Aristotle, she argues that despite its own emphatic claims to the contrary, medicine is not a science at all – and nor, incidentally, is it an art. Medicine is a practice – specifically, an uncertain, paradox-laden, judgement-dependent, science-using, technology-supported practice. As such, and despite the extensive pathophysiological and epidemiological evidence base that informs it, medicine is comparable to the practice of law or making of ethical judgements. In every case, the practitioner must reason not from the general to the particular but from the particular to the general – abduction rather than deduction. The question facing every practitioner, every time they encounter a case, is ‘What is it best to do, for this individual, at this time, given these circumstances?’

The good clinician must draw, as the founding fathers of evidence-based medicine famously pointed out, conscientiously and judiciously on the best research evidence on offer and make optimal use of available technologies (Sackett et al. 1996). But because the case is a singular one, because the predicament affects a real individual in the inescapable reality of the here and now and because this person has life projects and commitments and things at stake, a clinical judgement is fundamentally a practical and moral one, informed but not determined by scientific evidence. The skilled practice of medicine is not merely about knowing the rules but about deciding which rule is most relevant to the particular situation at hand. Illness may be a narrative, but just as in law, just as in literature, there is no text that is self-interpreting (Montgomery 2006).
As Montgomery points out, Aristotle distinguished three kinds of knowledge: *episteme* (factual knowledge), *techne* (technical knowledge or skill) and *phronesis* (practical wisdom). Clinical method involves not merely the skilled use of the senses to inform diagnosis but also the application of practical wisdom to draw judiciously on science (the evidence base) and technology (e.g. point-of-care prompts or inbuilt templates) when making here-and-now practical, ethical judgements in a particular case. Precisely how these different knowledges are woven together to manage and support the patient depends on the particularities of the case – and it occurred in different ways as I consulted with Lindsay, Chris, Aisha and Faisal.

**Clinical Method as Multi-professional Collaborative Work**

The input of a doctor to a twenty-first-century consultation usually both presupposes and contributes to a wider package of inter-professional care. When I saw Lindsay, she had already had a telephone consultation with a nurse practitioner and been advised to attend for an ‘on the day’ emergency appointment. When Chris registered with my practice a year ago, he had had an interview with a health care assistant, who had collected and entered a standard dataset of items such as weight, height, blood pressure and smoking status. More recently, the practice manager had contacted Chris to check missing data items so as to complete his Framingham risk score (a task which attracts an incentive payment for practices). These data items had been combined in an automated algorithm to produce a ‘high risk’ alert, which triggered the pop-up prompt in Chris’s electronic record at the point of care. Aisha’s record included an electronically transferred letter from the doctor in the orthopaedic clinic and the results of tests performed elsewhere in the health care system. Faisal, who attended with a professional interpreter (electronically booked on the system), had previously had an in-depth interview and physical examination by specialists in a third-sector organisation; a letter summarising their thorough, culturally informed assessment was on his electronic record when I saw him – but it was, sadly, technically inaccessible to me last Friday.

Whilst multi-professional care is part of the business as usual of both general practice and hospital care, few who write on clinical method make explicit reference to a wider team, each member of which typically contributes partially but not exclusively to the care of the patient or to the information systems which (well or badly) support this activity. Yet if we broaden our conception of clinical method to embrace the wider socio-technical care network (i.e. networks of people and the technological infrastructure which links them across time and space), the model of clinical care becomes an order of magnitude more complex. Uncertainty becomes plural.

There is my own uncertainty about the evidence base, about the unfinished aspects of the patient’s story and about the numerous interacting influences I need to take intuitive account of when making my here-and-now practical judgement of
what to do next. There is also my uncertainty about what others in the team know and what they have already done (and their uncertainties about what I have done).

Then there is the uncertainty inherent in data. Is this data item reliable? Is it complete? Is it up to date? Is it technically accessible? Do I trust the person who entered it? Does the absence of a data item indicate that the patient is ‘normal’ in relation to this item? These questions about data are not unique to the shared electronic record. They also pertain to shared paper records, patient-held shared care notes, patient-recorded data (e.g. home blood pressure or peak flow readings), post-it notes, telephone messages, emails and so on – but the distributed electronic record makes the trustworthiness and accessibility of data a particularly prominent issue. Finally, there is uncertainty about whether a prompt set up by some other person at some other time and in some other place is relevant to the patient sitting in front of me in the here and now (a phenomenon termed time-place distanciation (Giddens 1984)) and whether I should (sometimes or always) disrupt the flow of today’s consultation to respond to this interruption.

Given these complexities and interdependencies, it is perhaps small wonder that the leading cause of medical error is poor communication and collaboration between individuals and teams, including end-of-shift handovers, referral and discharge letters, real-time cross-referrals between members of acute teams, doctor-nurse communication in chronic disease management and transfer of prescriptions between physicians and pharmacists (Institute of Medicine 1999). It is often assumed that the solution to what has become known as the ‘integration’ problem is largely or wholly technological. In particular, many have speculated about a future electronic patient record which will be characterised by completeness, accuracy and extensive interoperability with other record systems (Institute of Medicine 2009). Data, it is assumed (or at least, implied), will be available at the touch of a button – at which point, uncertainty will become a thing of the past.

This framing of clinical care as wholly or mainly to do with data capture and retrieval suggests (wrongly) that knowledge about the patient can be accurately and completely recorded on the electronic record and passed as codified data items between professionals. ‘Continuity of care’ and ‘integration of care’ have become subtly redefined in technological rather than human terms. To explain why patients and staff alike so often experience multi-professional care in terms of discontinuity and lack of integration (and hence associated with greater rather than less uncertainty), we should remember the crucial role of tacit knowledge in clinical method:

Direct, face-to-face human interaction comprises a rich, highly nuanced exchange of information; the wealth of verbal and non-verbal communication it includes is necessarily absent from written records, telephone interactions and subsequent memories of an interaction. The full range of tacit and explicit information about a patient’s particular problem or illness is accessible only from within the clinical encounter. (Henry 2010, p. 293)

Technology-supported multi-professional care is a complex area of practice, the study of which owes much to an applied (and theoretically eclectic) field known as computer-supported cooperative work or CSCW (Berg 1998; Berg and Goorman 1999; Pratt et al. 2004). CSCW focuses not on the individual knower or computer
user but on the wider network of people and technologies and the work that is spread in complex ways among them. Inter-professional collaborative work involves the sequential ordering and coordination of tasks and the management of the inter-dependencies of these tasks, which in turn requires both real-time processing of local information by individuals and an awareness of how everyone’s contribution fits into the wider picture.

Shared electronic records tend to be seen as representing progress (e.g. greater accessibility, greater accuracy, searchability) from their paper-based predecessors. They can also (in theory) provide multiple views and framings of the data, hence can potentially tolerate (and overcome) the ambiguities inherent in inter-professional work and make the work of different professional groups more visible to others. In practice, however, this is rarely achieved, and the reality may be of ‘clunky’ interfaces, missing data and a sense of fragmentation and mutual alienation rather than inter-professional co-operation (Stange 2009). Despite this well-described model-practice gap, humans are often very creative in developing ways to get the job done (‘workarounds’), thereby overcoming inherent limitations of technologies and individual shortcomings of team members (Ash et al. 2004).

Some approaches to CSCW use the term ‘distributed cognition’ to depict knowledge and reasoning which are spread between people and technologies, perhaps across several departments or organisations. This model views people and computers as a linked set of information containers and processors, and implicitly views the key distinction between them as quantitative (e.g. computers have more memory and faster processing power). In this model, ‘uncertainty’ approximates to missing data items somewhere in the system.

More interesting for the purposes of this book are more nuanced approaches to CSCW whose starting point is the assumption that humans and computers are fundamentally different. The key accomplishment of the human team member is not in faithfully following a particular standardised routine (e.g. a shared protocol or computerized template) but in knowing the contingent detail of when (and when not) to follow that routine (Garfinkel 1967). This is what Garfinkel called ‘ethnomethodology’: ‘the moment by moment management of contingent detail through sequential orderings’ (Rawls 2008, p. 703). Only through effortful attention to situated detail can collaborative work occur, hence Garfinkel’s definition of a group as people who are, at any moment in time, ‘playing the same game’. A key dimension of this game-playing is trust, without which the continual production of meaningful, sequential action is impossible. The awkwardness between Faisal, myself and the interpreter (and in particular, her apparent interpretation of my silence as ignorance of which medication to prescribe for depression) illustrates how the three of us were playing ‘different games’ in this consultation.

Garfinkel influenced the seminal work of Lucy Suchman on situated action – that is, the subtle, contingent and context-dependent nature of action as humans use technologies collaboratively (Suchman 1987). Health care is a complex, nonroutine affair: contingencies are the rule, and the skill of professional practice is ‘smoothly molding such continuous lapses of order into events to be handled with “standard operating procedures”’ (Ash et al. 2004, p. 106). This essence of clinical work
contrasts markedly with the assumption inbuilt into the design of many health care information systems – that much work is routine and therefore readily automated.

All this has crucial (but often poorly understood) implications for the management of uncertainty in clinical care. First, a data item (such as a coded diagnosis of depression) may be present on the electronic record and appear factual but still have much uncertainty associated with it. Who entered this code? On what grounds? Is the diagnosis still current? What does the absence of a code for depression mean? What does three separate codes for myocardial infarction in a patient’s record mean – that the patient has had three heart attacks or that the quality control of data entry in the practice is poor?

Second, ‘overcomplete’ medical records may generate their own uncertainties through loss of overview (not seeing the wood for the trees) and information overload. Many of us remember as juniors sifting through dozens of pages of test results (paper or electronic) in a search for the one result that will change our management. In one extreme case, the introduction of a ‘cut and paste’ facility led to a junior doctor dumping an entire copy of a patient’s hospital admission notes into the electronic discharge summary (Ash et al. 2004). Aisha’s electronic record contained various imaging reports, some miscoded as ‘letter from consultant’, and numerous other items of correspondence, but no easily accessible overview of the extent or pace of deterioration over time.

Third, the collection and analysis of electronic record data for secondary uses such as audit, epidemiological research and surveillance of clinician performance is built on what is arguably an illusion – that it is only a matter of time before we will know exactly what is going on (e.g. what are the precise patterns of disease in this population? Who has been missed in each screening programme? Who are the ‘poorly performing’ general practitioners? and so on). In reality, there is a paradox: efforts to generate information (as in the Quality and Outcomes Framework in English general practice (Roland 2004)) may increase rather than decrease uncertainty. This is because (a) collecting and retrieving information both have an opportunity cost and divert activity from other, patient-facing work; (b) too much information leads to loss of overview at the policy level as well as the individual level; (c) incentivisation leads to gaming; and (d) information is rarely value neutral.

As Hari Tsoukas has put it in an article entitled *The Tyranny of Light* (Tsoukas 1997):

> The overabundance of information in late modernity makes the information society full of temptations. It tempts us into thinking that knowledge as information is objective and exists independently of human beings; that everything can be reduced into information; and that generating ever more amounts of information will increase the transparency of society and, thus, lead to the rational management of social problems. However… the information society is riddled with paradoxes that prevent it from satisfying the temptations it creates. More information may lead to less understanding; more information may undermine trust; and more information may make society less rationally governable. (p. 827)

The widespread notion that multi-professional care is best supported by the rationalist approach of standardised protocols, strictly delineated roles and an electronic record system that assures quality through coded entries, templates and
pull-down menus, has been challenged by academics working in the field of complex adaptive systems (Lanham et al. 2009). These authors argue that when uncertainty is high (i.e. most of the time in primary health care), quality is not something that is achieved through careful planning and adherence to protocol but something that emerges through adaptive relationships, collective sense-making and on-the-job learning from one another. The implications of this are far-reaching and beyond the scope of this chapter but deserve further exploration by those researching the link between teamwork and the management of uncertainty.

Conclusion

In this chapter, I have begun to develop a new taxonomy (which is no doubt incomplete, but it will do for a Friday evening surgery and some ideas which others can take forward). I summarise this taxonomy below.

First, there is uncertainty about the evidence – the ‘voice of medicine’ dimension of the consultation, for which the key questions relate to the completeness, accuracy and relevance of research-based evidence and on the balance between potential benefits and potential harms. Second, there is uncertainty about the patient’s story – the ‘voice of the lifeworld’ dimension, about which scholars of narrative medicine have offered much sound advice. Third, there is uncertainty about what best to do for a particular patient given a particular set of circumstances; this kind of uncertainty includes the philosophical question of how tacit knowledge informs clinical judgement. Finally, there are the many uncertainties (and associated threats to quality and safety) that inevitably arise when clinical care becomes a collaborative endeavour in which human-human, technology-human and technology-technology interactions all loom large.

Having artificially deconstructed uncertainty in clinical practice into four categories for analytic purposes, it is important to add that at a practical level, in the fast-moving and often down-and-dirty setting of front-line clinical work, uncertainty is a singular, shadowy and irrevocably fuzzy construct, not a multifaceted, tidy and well-defined one. Even when we try to be aware of uncertainty in the clinical consultation, it continually slips from our awareness.

It is for this reason, perhaps, that so many different groups from different countries (US, UK, Denmark, Sweden, Scotland, Canada) have independently arrived at a broadly similar way of addressing uncertainty in a general practice setting: the collective study of clinical practice through retrospective sense-making in groups. These examples differ substantially in the extent to which they view their work as grounded in particular academic traditions (e.g. in the psychodynamic focus of therapists such as Michel Balint (1957) and Carl Rogers (1951); the literary traditions of Michael Bakhtin (1981); the sociological perspective of Mike Bury and Arthur Frank (1995, 1998); the philosophical Aristotle’s practical reasoning and Nicomachean ethics taken forward by Katherine Montgomery and Rita Charon (Charon and Montello 2002; Montgomery 2006); and my own work on combining
evidence-based and narrative-based approaches to clinical decision-making (Greenhalgh 1999), as primarily educational (with a focus on defining learning objectives and measuring performance) or as a ‘branded’ approach with a focus on structured steps to be taken by a facilitator to support the group process (e.g. with a proper noun, as in Sommers’ Chapter on Practice Inquiry or Armson and MacVicar’s on Practice-Based Small Group Learning).

I suspect, however, that despite the differences in form and style, the commonalities between the approaches described in this book are more noteworthy than their differences. I am confident that committed participation in any one of these approaches will help clinicians in their struggle to do the best for their patients despite the inherent uncertainty of primary care practice. And I am also convinced that no matter how long such groups go on meeting or how much any specific approach is refined, there will never be a fix for those problems and situations that most trouble us. The most we can do with those is muddle through while we continue to reflect in supportive environments.

References


Clinical Uncertainty in Primary Care
The Challenge of Collaborative Engagement
Sommers, L.S.; Launer, J. (Eds.)
2013, XIV, 306 p. 41 illus., 18 illus. in color., Hardcover
ISBN: 978-1-4614-6811-0