

Development and significance of automated history-taking software for clinical medicine, clinical research and basic medical science

■ D. Zakim

From the Unit for Bioentrepreneurship (UBE), Medical Management Centre at the Department of Learning Informatics, Management and Ethics (LIME), Karolinska Institutet, Stockholm, Sweden

Keywords: medical history, computing, self-reported clinical data.

The best we can do is to divide all processes into those things which can be better done by machines and those which can be better done by humans.

John von Neumann, 1955

Introduction

The meaning of von Neumann's statement is that rule-based tasks are undertaken more faithfully by computers than by humans and that the potential benefit of computing can be assessed by evaluating whether there is less than acceptable adherence to applicable rules when a task is carried out by humans. von Neumann's prescient idea has been applied to essentially all theoretical, experimental and applied sciences and engineering. Today, computing is indispensable to these disciplines and fuels their progress. However, there has been little application of computing to the fundamental, rule-based processes underlying effective practice of medicine, specifically to collecting and interpreting clinical data from patients. The nature of information collected from examining patients is different, of course, from data in physical experiments. However, the processes for collecting these data, and their interpretation, are no less rule based. The absence of significant computing power in daily clinical practice cannot be attributed to lack of need as, whenever scrutinized, the processes by which physicians collect data from patients and make clinical decisions are noncompliant with applicable rules [1–10]. These failures contribute to less than acceptable outcomes for patients [11–14] and to as many as 98 000 deaths/year in the USA from medical errors [11]. Poor data collection in clinical practice also impedes progress in clinical [15–18] and basic science research [19, 20].

The purpose of this review is to summarize the development of automated history-taking software, considering the evidence for the validity and superiority of this method for collecting clinical data, and the continued need for history taking as the basis for quality care. Clinicians unfamiliar with programs for automated history taking need to begin to understand the power of the technology and its potential for improving clinical outcomes. Physicians are familiar with electronic health record (EHR) technology; however, EHRs do not incorporate formalized rules for history taking, cannot improve adherence to these rules and hence lack the benefits that come from automating rule-based tasks. True automation of history taking may be considered 'the second generation of EHRs' and, in our opinion, will be an inevitable development. This is the only mechanism for removing the incompatibility between unaided, finite human cognition and limited physician time and maximizing the value of healthcare at affordable cost.

Early innovations: 1950–2000

Automation of history taking and medical diagnosis

The knowledge base for medical practice is enormous, and is expanding continuously [21] and growing in complexity. Furthermore, established facts are subject to change. Apparently simple complaints can have extensive differential diagnoses. These cognitive loads are exacerbated by less time to see more patients. The profession's organized response to the cognitive challenges has been specialization so that each specialist manages repetitively a relatively narrow set of clinical issues. Specialization, however, has not improved clinical outcomes for patients, which are superior in geographical areas with higher proportions of primary

care physicians [22–25]. An alternative to specialization is the use of computing to reduce the cognitive load of practice. Efforts in this direction began in the late 1940s.

Brodman *et al.* [26] developed the Cornell Medical Index as a questionnaire to standardize data collection, orientated towards systems review to provide ‘... a quick and reliable method of obtaining important facts about a patient’s medical history without expenditure of the physician’s time’. It may seem trivial, but the Index demonstrated that history taking did not require a physician. In fact, the Index collected significantly more relevant information than physicians examining the same patients. Subsequently it was shown that a computer, when fed standardized, self-reported history data, was as accurate as physicians in diagnosing a wide array of common disorders [27]. This work thus dispelled the then prevalent notion that diagnosis was ‘... indefinable and intuitive...’ rather than ‘... logical and completely defined...’ [27]. The value of self-reported medical histories collected by questionnaire and processed by computer was replicated at Kaiser Permanente clinics using a questionnaire of 600 questions [28]. Results were reported only for the diagnosis of asthma. Sensitivity and specificity of the analytical method were good.

Slack and colleagues expanded the reach of self-reported history taking by programming a computer to collect the history directly from patients [29]. The rules for history taking were programmed by branching question trees so that the path from one question to another was determined by a prior answer and could be related to different, specific medical issues in different patients. This simple logic function made it possible to expand the number of questions posed by computer as compared with a paper questionnaire. This first computerized history-taking program was limited to allergy-related issues but was extended to separate programs for the review of systems and several problem-specific interviews [30]. The success of these programs was measured by the completeness of data collection as compared with physician-acquired histories, with which they compared favourably.

Mayne and colleagues at the Mayo Clinic were also early developers of computerized history-taking programs. Their program asked patients to select a chief complaint from a menu of complaints [31].

The program contained no follow-up to resolve the differential diagnosis of the chief complaint but encompassed a review of systems plus elements of the past medical history [31]. A program with clinical coverage similar to that developed at the Mayo Clinic was tested in a small number of inpatients by Grossman *et al.* [32] with similar results, i.e. far more clinical information was recorded by computers than by physicians.

Lawrence Weed was another pioneer in the field. Weed’s starting premise was different from that of his predecessors. He explicitly acknowledged that the knowledge for practice exceeded what anyone could learn and use for collecting relevant clinical data from patients and for interpreting the data to make accurate, timely diagnoses and that computing was a method for bridging this gap [33, 34]. His purpose was to improve decision-making through better data collection coupled to automated decision support. Weed’s history taking was committed to diagnosis by the closeness of fit between the set of clinical attributes displayed by a given patient with a specific chief complaint and the sets of attributes displayed by patients covering all the entities in the differential diagnosis of that chief complaint. For example, for a patient with vertigo/dizziness as the chief complaint, Weed’s scheme queried every patient for the presence of all attributes that could be displayed by any patient with any of the 78 diagnostic entities in his differential diagnosis for vertigo/dizziness [35]. The key output from the Weed program was not the facts of the history but the closeness of fit between the attributes of an interviewed patient and the attributes of a patient with each of the entities in the differential diagnosis of the patient’s chief complaint. Weed termed the history-taking routines ‘Problem Knowledge Couplers’. The ‘problem’ was the patient’s chief complaint, and the ‘knowledge’ was the database of attributes of well-characterized patients with known diagnoses. Published work with couplers does not cite a source for the ‘Knowledge’ databases. Furthermore, the overall scheme was not tested rigorously for utility and validity. A single clinical trial of the coupler program failed to establish diagnostic value as compared with routine care [36]. The patient population tested was relatively young, however, so this negative result does not mean that coupler programs are without value. Couplers have been reportedly used since 2005 only in a single small practice, but without data to assess how they are used and how they contribute to management [37].

The coupler programs are now owned by AskMD [38]. It seems unlikely that the couplers are being used as originally intended.

Many other investigators developed computerized history-taking programs during this period until 2000. A summary of these programs is provided in table 3 of the review by Bachman [39]. However, none of the programs cited expanded the technology beyond the range of issues addressed by Slack *et al.*, Mayne *et al.* and Weed.

Validity of self-reported clinical data in computerized histories

Self-reported history entries in the Cornell Medical Index questionnaire were documented by re-interview of patients and found to be highly accurate [26]. Discrepancies between self-reported entries for computer interview and physician records reflected equally low error rates by computer and physician [40]. Notably, however, computerized histories included more useful clinical data as compared with physician histories [40]. Analysis of data collected by self-reported, computerized history taking for pre-operative clearance showed that 97% of retest responses from 239 patients were identical to test answers [41]. A similar study of 48 patients interviewed using a general medical history program, showed a test-retest replication rate of 93% [42]. A quarter of the small number of discrepancies were patient errors in selecting the intended answers; another quarter occurred because patients were uncertain which of the available answers applied to them. Because self-reported clinical data acquired by computerized history taking are accurate, there is more clinically significant information in these histories as compared with medical charts. Indeed, comparisons of the accuracy of self-reported data and medical records [1, 2, 5, 7–9] show that the former are superior. This is not unexpected because computers do not interpret self-reported answers. There is no control over this aspect of history taking by physicians. Moreover, use of surrogates and scribes [43] in the era of EHRs makes entry of accurate clinical data into the record increasingly complex in ways avoided by automating the process.

Patient experience of self-reported computerized history taking

Even at a time when few had experience with or knowledge of computers, patients had a high level of acceptance of the technology for history taking

[29, 31, 44, 45]. This remains true now that computing is ubiquitous [42, 46]. Patient acceptance of computerized history taking extends beyond willingness to use these programs to belief that use will improve the quality of their healthcare [32]. Patients are also quick to humanize the computer [47, 48]. In addition, patients are willing to invest amounts of time inputting data that are not available to physicians, e.g. self-reported history-taking sessions exceeding 1 h [31, 42]. Another advantage of computerized history taking is that patients may reveal information about risky lifestyles to the computer that they will not share with physicians [49, 50]. It has not been tested for users of computerized history-taking programs; but studies of the impact of paper questionnaires on physicians show that patients using questionnaires are more satisfied with their care than those not providing the physician with self-reported clinical information [51].

Physician perception of clinical value from self-reported computerized histories

There has been no significant physician 'buy-in' to the use of computerized history-taking programs [31, 41]. Physicians who were questioned about the value of these programs believed that differences between their and computer findings were false-positive computer outputs or nonrelevant information [31, 41]. Quaak *et al.* [52] reported a similar view. Physicians who were not responsible for primary data collection were asked to compare the diagnostic value of information acquired by a computerized, general medicine history-taking program with information from physician-acquired histories [52, 53]. Although the computerized histories were acknowledged to contain more clinical information, physician histories were deemed more valuable for making diagnoses.

Given the test-retest reliability of automated histories, physician opinions of high rates of false-positive findings in self-reported histories are not correct. It seems instead that physicians consistently discount the value of collecting more clinical information than they typically acquire. This is likely to reflect the negative impact of limited human cognition on clinical decision-making. Thus, the information a physician can collect, analyse, record in a chart as an integrated note, and use to make a management decision is limited at all stages by a finite cognitive capacity. Physicians' opinions about the relevance of data to a

clinical decision depend on their cognitive capacity, with an upper limit susceptible to downregulation by distractions and time pressures. A computer has a limitless capacity for memory, recall and simultaneous factoring of very large numbers of variables not subject to downregulation by distraction and limited time. Hence, computerized history taking collects more clinically relevant information than a physician can process. The physician will regard some of this 'excess' as superfluous to decision-making especially when the information is unrelated to the acute problem or encompasses unfamiliar aspects of medicine. This mismatch between what computers can output and what physicians can make sense of is already a common problem. EHR alerts [54], for example, are '... rendered meaningless by their sheer number' [55]. The value of automated history taking cannot be determined, therefore, by physician opinions about the data collected by computer but must depend on the impact of the information on outcomes for patients.

Computerized history taking might collect and output information that a careful physician will detect as trivial and incorrect. But well-designed software follows positive findings to determine their relevance to a specific differential diagnosis and in this way finds inconsistencies as they occur during an interview. For example, when an entry violates a programmed rule, the program can pose a question to clarify inconsistencies for both false-positive and false-negative entries. Another method for improving automated history taking is to reduce the cognitive load of reporting to physicians by organizing findings as pathophysiological narratives. This method facilitates comprehension of complex information compared with a list of facts. Inadequate attention to the problem of conveying complex information probably underlies, in part, the negative results cited by Quaak *et al.* [52]. It is difficult to grasp the meaning of data in the output from their program [56].

Development in the 21st century

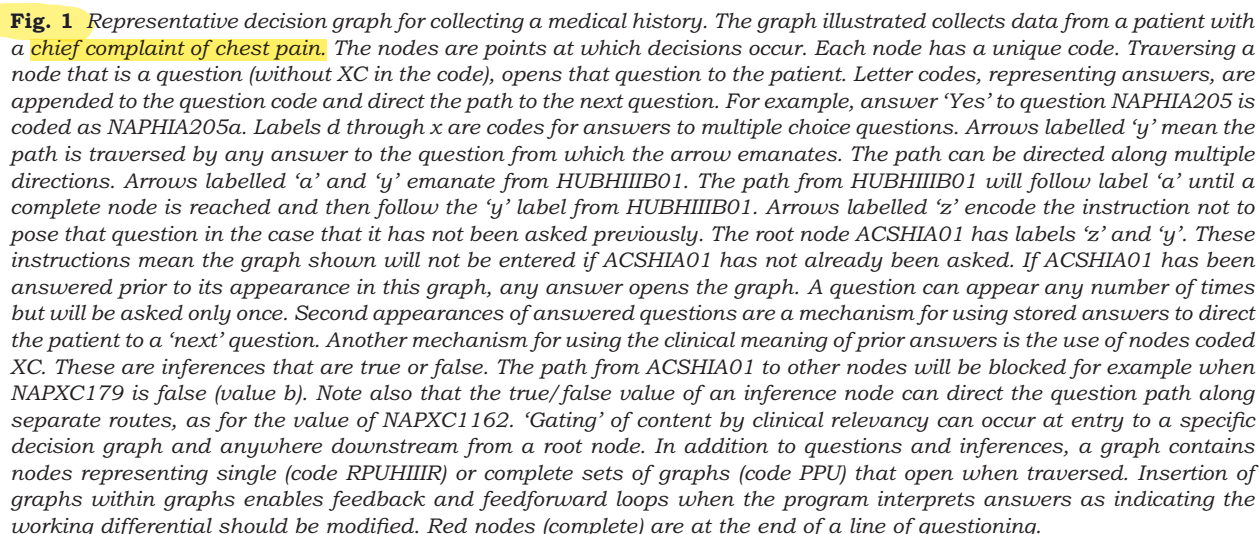
Specialized history-taking programs

Some of the programs cited by Bachman [39] may be in use at institutions in which they were developed, but none seems to be in use elsewhere except for Instant Medical History [57, 58]. History-taking programs developed more recently by the academic community [59–62] interact with patients on single issues and seem to have had

limited use outside their sites of development. An exception is web-based, self-reported assessment of disease activity by patients with rheumatoid arthritis in Sweden [63]. In addition, The Science of Your Cycle [64] is a clinically directed automated history-taking application to track a woman's menstrual cycle, '... the next period, fertile window and PMS' and '... physiological processes ... throughout the cycle'. The Science of Your Cycle appears to deliver personalized advice to affect clinical outcomes by using data-driven analytical algorithms to personalize prediction and to leverage the power of standardized data to improve its analytics. There is no comparable program within the academic medical community. This history-taking program indicates how the future of practice may be impacted by nonmedical communities using computer technology to meet limited goals.

Complete history-taking programs

Slack *et al.* [42] have expanded on earlier work to create a general medicine, history-taking program. However, no details about its operation have been reported; nor is it clear that the program is in use. Instant Medical History is advertised as a method by which the patient's answers populate data fields in the physician's EHR [58]. Details of coverage and operation are not available. The program CLEOS® (Clinical Expert Operating System) may be the only automated history-taking program under continuous development for acquiring complete medical histories [65]. CLEOS® differs from other history-taking software described in the literature in that it emulates clinical thinking continuously as data are collected. The key to this is automated interpretation of all prior answers as the interview proceeds so that the next questions are determined by pathophysiology that relates all prior answers to the working differential diagnosis. This mechanism also enables continuous checking for entry errors and inconsistent answers. Fig. 1 is an example of the decision graphs that determine the pathway of questioning in a CLEOS® interview. How the logic functions operate to determine a pathway through the graph is described in the legend. Fig. 2 is a graphic representation of the more than 450 component decision graphs of the complete program at present. Every patient enters Fig. 2 by selecting a chief complaint at the point marked by the arrow and traverses the complete pathway of the graph of about 17 000 decision nodes. The logic functions (Fig. 1) determine which questions are relevant to



questions not asked are nonrelevant to the patient's medical issues. There is no limit to the number of graphs that can be accommodated in

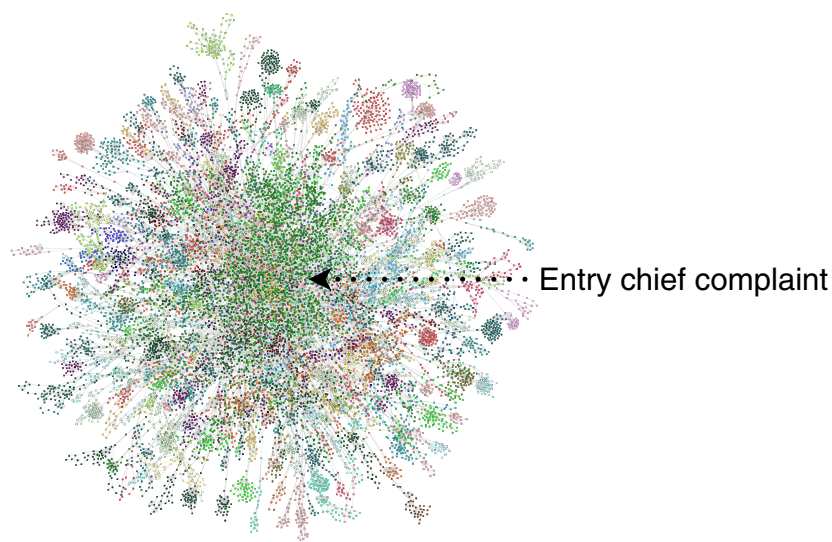


Fig. 2 Graphic representation of the complete CLEOS® program. The complete program comprises approximately 450 different decision graphs as in Fig. 1. The arrowhead indicates the entry point for a first history-taking session with entry of a chief complaint. Each dot is a decision node with codes and significance for guiding the interview pathway as in Fig. 1.

Fig. 2 and the range of clinical and lifestyle issues the program can address. With regard to the output, CLEOS® uses analytics to report findings as pathophysiological narratives, differential diagnoses, management recommendations and orders for diagnostic and confirmatory laboratory testing and suggested treatment plans [46, 65]. This is another important difference between CLEOS® and other automated programs, which report only factual findings. CLEOS® out-performs physicians using the metric of significant clinical problems identified [52], which is a direct measure of differences between computer and physician in finding clinically significant information. CLEOS® also out-performs physicians in adherence to guidelines, which was measured for the management of LDL cholesterol [66].

CLEOS® stores data elements as codes. This facilitates rule-writing (Fig. 3) and mathematical analyses of standardized data sets [67]. Rules that are true for the data in a given interview are also stored as codes. These can be managed by computer to improve methods for conveying complex clinical information. Furthermore, because clinical management is seldom a one-off event, automated history-taking programs have to function across time to provide automated data collection relevant to follow-up. This capability depends on internal intelligence to identify problems in the course of disease and clinical attributes expected to change in the process towards resolution or guideline-compliant achievement of targets. CLEOS® supports automated follow-up in the context of

identified problems and can incorporate physical examination and laboratory results for this purpose. CLEOS® is owned by a Swedish nonprofit foundation for which the governing board is appointed by the President, Karolinska Institutet, Stockholm, Sweden. Knowledge content and access to databases generated from the use of this knowledge are controlled by the academic medical community.

Incorporating computerized, self-reported history taking into routine practice

Computerized history taking transfers the time and cognitive demand for collecting clinical phenotype data from the physician to the patient and a computer. The technology can populate whichever data fields may be needed for automated decision support across a wide range of clinical issues. Only a single copy of the software needs to incorporate new knowledge to enable rapid translation of validated clinical advances to everyday practice. Additionally, by using a machine, the collection of clinical data can be standardized across all patients in contact with the healthcare system and included in a database to fuel clinical and basic medical research. With regard to development, it is possible to formalize the entirety of medical knowledge on a hard drive that can be purchased for a few hundred dollars. There are tools with which anyone who can use a Word program can formalize medical knowledge as software. Computing devices with readable screens are ubiquitous. Problems in delivering affordable care

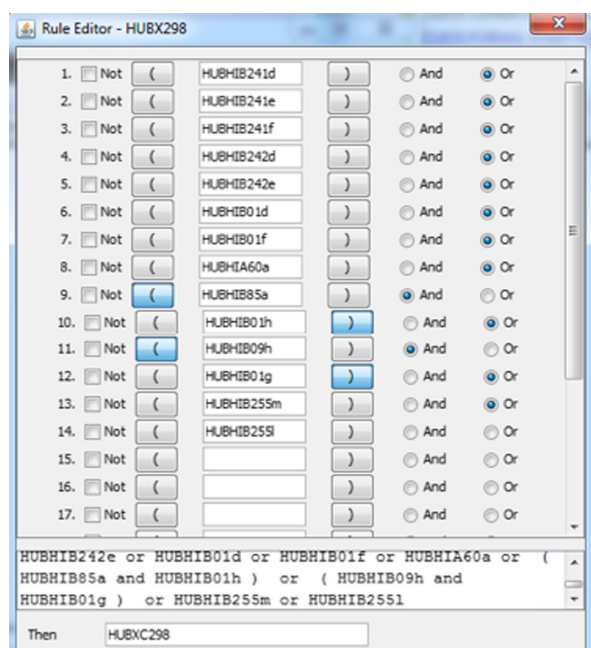


Fig. 3 Interface for writing rules in which arguments are answer codes, as shown here, or codes for another rule. The code for the rule shown is HUBXC298. Codes for rules have no appended lower case letters. The statement 'If HUBXC298...' has the implicit meaning that HUBXC298 is true. Arguments in a single rule can include answer codes, physical examination, laboratory result codes and codes for another rule. Arguments are connected by Boolean operators 'AND', 'OR' and 'NOT'. Parentheses are allowed but only as illustrated in this example

remain. Yet, after 50 years of development and validation, automated history taking is essentially not used in medical practice. Available evidence suggests that physicians who developed these programs focused more on demonstrating what was possible rather than on building sustainable programs that would be adopted widely by the clinical community [68]. Achieving widespread use depends on attracting young physicians to the programme for developing automated history taking, and maintaining continuous expansion and refinement of the medical knowledge incorporated into the software, continuous thinking to maximize the value of computing power and a rigorous schedule of clinical testing. The clinical community and patient advocacy groups must be engaged in these processes. In addition, testing for clinical efficacy must establish that automated history taking can impact the only goal that matters: clinical outcomes for patients. To our knowledge,

only one trial has been conducted to date that measured an effect on outcome. In this trial, selected established patients in primary care interacted with Instant Medical History from home [69]. Physicians determined from these data, available electronically, that 40% of 2500 patients did not need care. Their medical issues and how they differed from patients deemed to need office care were not provided. However, this is the only clinical outcome reported so far to be affected by the use of automated history taking.

Bringing automated history taking to everyday clinical use remains a sizable undertaking requiring an interdisciplinary team of specialist and primary care physicians, patient advocates, computer science, big-data analytics and experience design. A significant start has been made in recruiting this expertise by the development of CLEOS[®] within the Karolinska Institutet. We foresee further development and maintenance of CLEOS[®] as a cooperative, worldwide effort within the medical community, and are unaware of similar developments at any other institution.

Is the medical history still relevant?

Most physicians will subscribe to Osler's advice: 'Listen to the patient. He is telling you the diagnosis' [70]. Medical educators affirm this guidance, and text book discussions of the pathophysiology of a disease and its management begin with the instruction to take a history. In contrast to these 'official' positions, 'bedside' skills are considered less important than in the past [71–73] and performed poorly when residents are observed directly [74, 75]. Substantial numbers of young internists feel poorly prepared for interviewing patients [76]. Physicians are substituting laboratory data for history taking [77–82]. In addition, it is believed that medical practice is about to be revolutionized by genomics and related fields that will provide all the information needed to predict risks, make diagnoses and personalize treatment decisions. Given the difficulty in taking a history of sufficient detail to maximize the value of medical knowledge for patient outcomes, there would be enormous benefit from laboratory tests providing information at least equal in value to the detailed examination of patients. When tested in specific settings, however, the superiority of standard laboratory tests and/or imaging as compared with examining the patient cannot be substantiated [83–85]. More generally, there is an inverse

correlation between the use of laboratory testing and quality of outcomes for patients [22, 23, 82, 85]. In addition, although genetic markers appear to be useful for managing selected tumours [86–88], and this application is likely to expand, there is no experimental support for a soon-to-occur, genomics-driven, revolution in medical practice. Genomic and proteomic measurements provide information on the parts of cells. Normal function and disease result from interactions and disarray of interactions, respectively, between these parts, between complex systems in cells, and between cells, tissues and organs [89]. Of course, interacting molecules can be studied *in vitro*; however, it is extremely difficult *in vitro* to elucidate and account for the complexity of molecular interactions in cells of interest [90, 91]. A structural change in a specific 'part' can have variable functional consequences for different interacting systems of which it is a component. The effect on function from a structural change in any given part can depend on the structures of all other parts of the interacting system, i.e. epistasis is pervasive. Also, studies of two or a few interacting molecules *in vitro* cannot model interactions between systems, organization of cells into organs, interconnectedness of organ functions and the influence of the external environment on these complexities. In other words, functions of intact cells, organs and animals are too difficult to model meaningfully from molecular data [89, 92–96]: '... there is no way to gather all the relevant data about each interaction included in the model' [90, 97, 98]. In contrast with molecular level information, information from clinical examination reports on the overall status of complex interactions, i.e. whether normal or dysfunctional as well as the array of normal and dysfunctional interactions in an individual and how these interactions change with time. The best evidence, therefore, is history taking still provides the key data set for decision-making to maximize the quality of care for patients and that this will not change for the foreseeable future. Indeed, the most significant change in practice since Osler's time is the enormous expansion of the knowledge that informs practice, which, for effective use, requires more information about the patient than in the past.

Integration of clinical research into clinical practice

History taking is the only aspect of clinical data collection not studied from the perspective of how to maximize value. Automation of history taking in

everyday practice makes it possible to study this critical component by determining what information needs to be collected in which patients. Consider the example of prescribing statins to prevent ischaemic events. Applying the current guideline will prevent 1–2 events per 100 persons treated [99, 100]. This result means that outcomes in statin trials [101] are affected by uncontrolled variables [102] or that risk of ischaemic events cannot be determined for individuals with odds better than about 50 : 1. We can test whether this is a true upper limit on knowledge for predicting ischaemic events as a routine feature of everyday practice by employing automated history taking. Tables 1 and 2 show variables likely to influence ischaemic risk that are not controlled in statin trials and that are not included in statin guidelines [103]. The variables are labels for complex processes. For example, 'Past medications, Stopped for nonefficacy' (Table 2) provides information, medication by medication, on a linked set of complex processes that cannot be modelled *in vitro* [104]. In addition, basic biology indicates that the complex systems on which the labels report are unlikely to function independently of each other. There are millions of persons at risk of ischaemic attacks, many taking statins prescribed haphazardly [65, 105]. Automated history taking can acquire data shown in Tables 1 and 2 without expending physician time and store the data for analysis without burdening physicians with clinical facts seemingly irrelevant to decision-making at the present time. Data mining '... to find unsuspected relationships and to summarize the data in novel ways ...' [101] can be applied to the very large, complex data sets that will be collected. Outcomes, for example ischaemic events, can be mapped to subgroups determined by data mining techniques. Should these analyses improve prediction on an individual basis, minimal data sets to achieve this goal can be defined, and new guidelines, together with the data elements supporting a clinical decision to treat, can be delivered automatically to physicians on a patient-by-patient basis. This general scheme is applicable to any problem for which there are no or inadequate guidelines. Thus clinical research can become an integral component of routine care [106–109]. Automated history taking can connect every practice and everyone receiving healthcare to the framework of clinical research. Automated history-taking also will have a positive effect on the progress of basic medical research [19, 20]. The key to the quality of clinical data for correlation

Table 1 Factors that can affect risk of ischaemic events but are not considered for calculating risk scores

Data set	Data elements
Demographic	Ethnic origin, place of birth; education; economic status; current occupation; occupational history
Habitus	BMI; waist/hip ratio; lifetime weight changes
Environment	Physical activity, current and lifetime; diet, current and history; current residence by postcode; potential toxin exposure Use of tobacco: type, quantity, duration; exposure to secondary smoke Use of alcohol, current and history Use of cocaine, current and history All current medications All past medications with ADEs
Epigenetic	Birth weight (or gestational age at birth), birth order, maternal diabetes during pregnancy
Comorbidities with duration, severity, response to treatment	Presence/absence of chronic inflammatory diseases; renal disease; pulmonary disease; or noncoronary heart disease Cancer history and treatment with cardiotoxic agents

Except for smoking, data elements listed are not considered in evaluating risk of ischaemic events. Smoking history in risk calculations does not consider individual differences in amounts and types of tobacco consumed. Nonsmoking status in risk calculations does not discriminate between never smoked and stopped smoking nor the duration of smoking in past smokers and interval since last smoked. Exposure to second-hand smoke in nonsmokers is not factored into risk scores. ADE, adverse drug events.

with findings at the molecular level is the scope of the data and its standardization across very large numbers of individuals.

There is a current view that EHRs will have an impact on clinical and basic science research similar to that envisioned for automated history taking [110, 111]. However EHR data cannot support a robust programme of clinical research. Because of the greater time demand of EHR use compared with hand-written or dictated notes [112], data entered into EHRs may be collected by physicians, by personnel with nonmedical training [113] or by scribes [43]. Physicians using EHRs do not collect data by a standard protocol or even with a standardized EHR programme. Furthermore, the integrity of data in EHRs is suspect for many other reasons [114–117].

Effect of automated history taking on physicians and physician–patient interactions

The physician reading the output from an automated history-taking program is in the position of a consultant providing a second opinion. This will enhance interactions with patients by removing the burden of data collection and thereby providing time to discuss in more detail the patient's issues

and options. In addition, automated history taking will avoid the use of current EHRs, which can interfere with patient interaction [118]. Physicians have embraced computer-based technologies across medicine, including imaging technology, Lasik eye surgery, robotic surgery in general and auto-analysers for laboratory medicine. However in these uses, computer technology does not provide patients with insight into the processes of practice. History-taking computers provide patients with a view of a key process by which medicine is practised that they might even understand. Physician resistance to use of automated history-taking in prior work may be rooted partly in this issue [22, 41, 52]. Since the power of automated history-taking is necessary for achieving better outcomes at affordable cost [11, 12], this question needs to be studied by involving clinicians in discussions and development of the technology as it emerges. It will also be important to start acknowledging that physicians, like everyone else, have finite cognitive abilities and that medical practice is an extraordinarily difficult cognitive task. Additionally, it may be important to note the analogies between robotic surgery and computerized history taking. Surgical robots provide a better view of the anatomy and surgical problem confronted than the unaided eye. The history-taking robot provides a more precise

Table 2 Factors related to the presence of diabetes and/or hypertension that affect incidence of complications but that are not considered in estimating risk of coronary disease

Comorbidity	Characterization of comorbidity
Diabetes	Age at onset
	Method of discovery
	Regulation of blood sugar:
	HbA1c, daily variation of blood glucose
	Hypoglycaemic events
	Severity, frequency, manifestations
	Current hypoglycaemic agents
	Adherence
	Past hypoglycaemic agents
	Stopped due to ADE
	Stopped due to nonefficacy
	Reversible vision changes
	Nonvascular complications (e.g. neuropathy, gastropathy)
	Vascular complications other than coronary
Hypertension	Age at onset
	BP at diagnosis
	BP current
	Self-monitored BP
	Current medications
	Adherence
	Past medications
	Stopped due to ADE
	Stopped due to nonefficacy
	Vascular complications other than coronary

Diabetes and hypertension are variables included in vascular risk scores. But the variable clinical characteristics of these states in different patients are not factored into risk scores. Treatment details for both diseases specify adverse drug events (ADEs) and agents with no efficacy because this information reports on complex interactions that are otherwise not available in any other data set (see [106]). HbA1c, haemoglobin A1c; BP, blood pressure.

view of the anatomy of problem(s) than is otherwise available to the unaided clinician. Problem(s) revealed are understandable only to a trained physician. What is missing in the analogy is evidence for better clinical outcomes from the use of history-taking robots. This is crucial for successful employment of automated history-taking programs.

Conclusion

Data from direct history taking and examination of patients remains essential for delivering quality outcomes. Contrary with expectation, advances in genomics and related fields will not revolutionize data collection in everyday practice. Indeed, the range of information available from interview of patients that should be factored into clinical decisions is expanding. Furthermore, in addition to the cognitive demand of history taking, there is less time for data collection. These critical problems in routine practice can be addressed with computer technology because history taking is a rule-based task. It is established that computers can acquire medical histories by direct interview of patients, that automated history taking collects more clinically significant information than physicians examining the same patients and that the data are valid. Lay persons have a positive response to interview by computer. At present, software is sufficiently powerful to extend automated data collection to any domain of interest, to emulate detailed clinical thinking and to enable computers to be critical observers of and interactors with the patient. Automated, self-reported history taking is a key technology for maximizing outcomes for patients, and will improve the physician–patient dialogue. It is also increasingly clear that clinical practice and clinical research must be integrated to improve the validated knowledge base for practice, which can only be achieved through automated history taking.

Conflict of interest statement

DZ declares no conflicts of interest. DZ is the inventor on US patents for technology related to the CLEOS® program that are assigned without any royalty rights to Stiftelsen CLEOS® Foundation in Stockholm.

Acknowledgements

Studies related to the development of CLEOS® have been supported in part by the Robert Bosch Stiftung, Stuttgart, Germany. The author thanks Dr. Andy Noori for programming support in developing CLEOS® and innumerable colleagues at the Robert Bosch Krankenhaus, Stuttgart for help in testing the program, with special thanks to Profs Mark Dominik Alscher and Matthias Schwab. The author is indebted to Prof Carl Johan Sundberg, Karolinska Institutet, for his leadership of the

CLEOS® program at Karolinska Institutet and for his careful and helpful reading of the manuscript.

References

- Bentsen BG. The accuracy of recording patient problems in family practice. *J Med Educ* 1976; **51**: 311–6.
- Romm FJ, Putnam SM. The validity of the medical record. *Med Care* 1981; **19**: 310–5.
- Beckman HB, Frankel RM. The effect of physician behavior on the collection of data. *Ann Int Med* 1984; **101**: 692–6.
- Ramsey PG, Curtis JR, Paaauw DS, Carline JD, Weinrich MD. History- taking and preventive medicine skills among primary care physicians: an assessment using standardized patients. *Am J Med* 1998; **104**: 152–8.
- LeBlanc VR, Brooks LR, Norman GR. Believing is seeing: the influence of a diagnostic hypothesis on the interpretation of clinical features. *Acad Med* 2002; **77**: S67–9.
- Cox JL, Zitner D, Courtney KD, et al. Undocumented patient information: an impediment to quality of care. *Am J Med* 2003; **114**: 211–6.
- Farmer SA, Roter DL, Higgenson IJ. Chest pain: communication of symptoms and history in a London emergency department. *Pat Ed Counsel* 2006; **63**: 138–44.
- Burnett SJ, Deelchand V, Franklin BD, Moorthy K, Vincent C. Missing clinical information in NHS hospital outpatient clinics: prevalence, causes and effects on patient care. *BMC Health Services Res* 2011; **11**: 14.
- Eze-Niliam C, Cain K, Bond K et al. Discrepancies between the medical record and the reports of patients with acute coronary syndrome regarding important aspects of the medical history. *BMC Health Services Res* 2012; **12**: 78.
- Institute of Medicine. National Academies of Sciences, Engineering, and Medicine. Improving diagnosis in health care. Washington, DC: (<http://iom.nationalacademies.org/Reports/2015/Improving-Diagnosis-in-Healthcare.aspx>). 2015.
- Kohn LT, Corrigan JM, Donaldson MS, eds. *To Err is Human*. Washington, DC: National Academies Press, 1999.
- Committee on Quality of Health Care in America. *Institute of Medicine: Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academies Press, 2001.
- McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, Kerr EA. The quality of healthcare delivered to adults in the United States. *NEJM* 2003; **348**: 2635–45.
- Wolff JL, Starfield B, Anderson G. Prevalence, expenditures, and complications of multiple chronic conditions in the elderly. *Arch Int Med* 2002; **162**: 2269–76.
- Ioannidis JPA. Why most published research findings are false. *PLoS Med* 2005; **2**: e124.
- Ioannidis JPA. Contradicted and initially stronger effects in highly cited clinical research. *JAMA* 2005; **294**: 218–28.
- Little RJ, D'Agostino R, Cohen ML et al. The prevention and treatment of missing data in clinical trials. *NEJM* 2012; **367**: 1355–60.
- Zakim D, Schwab M. Data collection as a barrier to personalized medicine. *TIPS* 2015; **36**: 68–71.
- Freimer N, Sabatti C. The human phenome project. *Nat Genet* 2003; **34**: 15–21.
- Riggs ER, Jackson L, Miller DJ, Van Vooren S. Phenotypic information in genomic variant databases enhances clinical care and research: the international standards for cytogenomic arrays consortium experience. *Human Mut* 2012; **33**: 787–96.
- Bastian H, Glasziou P, Chalmers L. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med* 2010; **7**: e1000326. doi:10.1371/journal.pmed.1000326.
- Starfield B, Shi L, Grover A, Macinko J. The effects of specialist supply on populations' health: assessing the evidence. <http://content.healthaffairs.org/cgi/content/abstract/hlthaff.w5.97.v1>.
- Baicker K, Chandra A. The productivity of physician specialization: evidence from the medicare program. *AEA Papers Proc* 2004; **94**: 357–61.
- Cooper RA. States with more physicians have better-quality health care. *Health Aff* 2009; **28**: w91–102.
- Baicker K, Chandra A. Cooper's analysis is incorrect. *Health Aff* 2009; **28**: w116–8.
- Brodman K, Erdman AJ, Lorge I, Wolff HG, Broadbent TH. The Cornell medical index an adjunct to medical interview. *JAMA* 1949; **140**: 530–4.
- Brodman K, Van Woerkom AN, Erdman AJ, Goldstein LS. Interpretation of symptoms with a data-processing machine. *Arch Int Med* 1959; **103**: 776–84.
- Collen MF, Rubin L, Neyman J, Dantzig GB, Baer RM, Siegel AB. Automated multiphasic screening and diagnosis. *Am J Public Health Nations Health* 1964; **54**: 741–50.
- Slack WV, Hicks GP, Reed CE, Van Cura LJ. A computer-based medical-history system. *NEJM* 1966; **274**: 194–8.
- Mayne JG, Weksel W, Sholtz PN. Toward automating the medical history. *Mayo Clin Proc* 1968; **43**: 1–25.
- Slack WV. A history of computerized medical interviews. *MD Comput* 1984; **1**: 52–9.
- Grossman JH, Barnett O, McGuire MT, Swedlow DB. Evaluation of computer-acquired patient histories. *JAMA* 1971; **215**: 1286–91.
- Weed LL. The computer as a new basis for analytic clinical practice: coupling of individual problems with medical knowledge. *Mt Sinai J Med* 1985; **52**: 94–8.
- Weed LL. New connections between medical knowledge and patient care. *BMJ* 1997; **315**: 231.
- Weed LL, Weed CC. The database: its definition and present status. In: Weed LL, ed. *Knowledge Coupling*. New York: Springer, 1991.
- Apkon M, Mattera JA, Lin Z et al. A randomized outpatient trial of a decision-support information technology tool. *Arch Int Med* 2005; **165**: 2388–94.
- Burger C. The use of problem-knowledge couplers in a primary care practice. *Permanente J* 2010; **14**: 47–50.
- <https://askmd.sharecare.com/askmd/all-consultations>
- Bachman JW. The patient-computer interview: a neglected tool that can aid the clinician. *Mayo Clin Proc* 2003; **78**: 67–78.
- Carr AC, Ghosh A, Ancill RJ. Can a computer take a psychiatric history? *Psychol Med* 1983; **12**: 151–8.
- Lutner RE, Rozen MF, Stocking CB et al. The automated interview versus the personal interview. *Anesthesiology* 1991; **75**: 394–400.
- Slack WV, Kowaloff HB, Davis RB et al. Evaluation of computerized medical histories taken by patients at home. *JAMA* 2012; **19**: 545–8.

- 43 Gellert GA, Ramirez R, Webster SL. The rise of the medical scribe industry implications for the advancement of electronic health records. *JAMA* 2014; **313**: 1315–6.
- 44 Quaak MJ, Westerman RF, Schouten JA, Hasman A, Bemmél JH. Patient appreciations of computerized interviews. *Med Inform* 1986; **11**: 339–50.
- 45 Slack WV, van Cura LJ. Patient reaction to computer based medical interviewing. *Comput Biomed Res* 1968; **1**: 527–31.
- 46 Zakim D, Braun N, Fritz P, Alscher MD. Underutilization of information and knowledge in everyday medical practice: evaluation of a computer-based solution. *BMC Med Inform Decis Mak* 2008; **8**: 50.
- 47 Quaak MJ, Westerman RF, Schouten JA, Hasman A, van Bemmél JH. Appraisal of computerized medical histories: comparisons between computerised and conventional records. *Comp Biomed Res* 1986; **19**: 551–64.
- 48 Zakim D, Alscher MD, Schwab M, Scwalm B, Sundberg CJ. Automated history-taking software: design, deployment and patient perceptions (in Swedish). *Läkartidningen* 2014; **111**: 2–5.
- 49 Erdman HP, Klein M, Greist JH. Direct patient computer interviewing. *J Consult Clin Psych* 1985; **53**: 760–3.
- 50 Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL. Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. *Science* 1998; **280**: 867–73.
- 51 Hershey CO, Grant BJB. Controlled trial of a patient-completed history questionnaire: effects on quality of documentation and patient and physician satisfaction. *Am J Med Quality* 2002; **17**: 126–34.
- 52 Quaak MJ, Westerman RF, van Bemmél JH. Comparisons between written and computerised histories. *BMJ (Clin Res Ed)* 1987; **295**: 184.
- 53 Peiris RD, Gregor P, Alm N. The effects of simulating human conversational style in a computer-based interview. *Interact Comput* 2000; **12**: 635–50.
- 54 Singh H, Spitzmueller C, Petersen NJ, Sawhney MK, Sittig DF. Information overload and missed test results in electronic health record-based settings. *JAMA Intern Med* 2013; **173**: 702–4.
- 55 Rosenbaum L. Transitional chaos or enduring harm? The EHR and the disruption of medicine. *NEJM* 2015; **373**: 1585–6.
- 56 Quaak MJ, Van Der Voort PJG, Westerman FR, Hasman A, Van Bemmél JH. Automation of the patient history – evaluation of ergonomic aspects. *Int J Bio-Med Comput* 1987; **2**: 287–98.
- 57 Wenner AR, Ferrante M, Belser D. Instant medical history. *Proc Annu Symp Comput Appl Med Care* 1994; **1036**.
- 58 <http://www.medicalhistory.com/home/index.asp>.
- 59 Vodstrel LA, Hocking JS, Cummings R *et al*. Computer assisted self interviewing in a sexual health clinic as part of routine clinical care; impact on service and patient and clinician views. *PLoS ONE* 2011; **6**: e18456.
- 60 Herrick DB, Nakhasi A, Nelson B *et al*. Usability characteristics of self-administered computer assisted interviewing in the emergency department: factors affecting ease of use, efficiency, and entry error. *Appl Clin Inform* 2013; **19**: 276–92.
- 61 Edelman EA, Lin BK, Doksum T *et al*. Implementation of an electronic genomic and family health history tool in primary prenatal care. *Am J Med Genet* 2014; **166C**: 34–44.
- 62 Almario CV, Chey W, Kaung A *et al*. Computer-generated vs physician documented history of present illness (HPI): results of a blinded comparison. *Am J Gastro* 2015; **110**: 170–9.
- 63 Ovetreit J, Keller C, Hvitfeldt Forsberg H, Essén A, Lindblad S, Brommels M. Continuous innovation: developing and using a clinical database with new technology for patient-centred care-the case of the Swedish quality register for arthritis. *Int J Qual Health Care* 2013; **2**: 118–24.
- 64 Tabor VH, Teel K, Wheeler MV, LaVigne M. The science of your cycle: evidence-based app design. <https://medium.com/clued-in/the-science-of-your-cycle-evidence-based-app-design-1969eab3fbc6#nz2b1h8vi>.
- 65 Zakim D, Alscher MD. A comprehensive approach to the IT: clinical practice interface. In: Wickramasinghe N, Bali RK, Kirn S, Suomi R. eds. *Critical Issues for the Development of Sustainable Ehealth Solutions*. New York: Springer, 2012; 353–74.
- 66 Zakim D, Fritz C, Braun N, Fritz P, Alscher MD. Computerized history-taking as a tool to manage LDL-cholesterol. *Vasc Health Risk Manage* 2010; **6**: 1039–46.
- 67 Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning Data Mining, Inference, and Prediction*. New York: Springer, 2009.
- 68 Barnett GO. Computers in patient care. *NEJM* 1968; **279**: 1321–7.
- 69 Adamson SC, Bachman JW. Pilot study of providing online care in a primary care setting. *Mayo Clin Proc* 2010; **85**: 704–10.
- 70 <http://www.osler.co.uk/william-osler/>
- 71 Feddock CA. The lost art of clinical skills. *Am J Med* 2007; **120**: 374–8.
- 72 Engel GL. Are medical schools neglecting clinical skills? *JAMA* 1976; **236**: 861–3.
- 73 Bordage G. Where are the history and physical? *Can Med Assoc J* 1995; **152**: 1595–8.
- 74 Sharma SA. Single-blinded, direct observational study of PGY-1 interns and PGY-2 residents in evaluating their history-taking and physical examination skills. *Perm J* 2011; **15**: 23–9.
- 75 Holmboe ES, Hawkins RE, Huot SJ. Effects of training in direct observation of medical residents' clinical competence: a randomized trial. *Ann Intern Med* 2004; **140**: 874–81.
- 76 Mueller PS, Barrier PA, Call TG *et al*. Views of new internal medicine faculty of their preparedness and competence in physician-patient communication. *BMC Med Educ* 2006; **6**: 30.
- 77 Wang CS, Fitzgerald JM, Schulzer M, Mak E, Ayas NT. Does this dyspneic patient in the emergency department have congestive heart failure? *JAMA* 2005; **294**: 1944–56.
- 78 Swap CJ, Nagurney JT. Value and limitations of chest pain history in the evaluation of patients with suspected acute coronary syndromes. *JAMA* 2005; **294**: 2623–9.
- 79 Statland BE, Winkel P. Utilization review and management of laboratory testing in the ambulatory setting. *Med Clin N Am* 1987; **71**: 719–32.
- 80 Alexander B. Reducing healthcare costs through appropriate test utilization. *Crit Values* 2012; **5**: 6–8.
- 81 Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaout R. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. *PLoS ONE* 2013; **8**: e78962.
- 82 McCormick D, Bor DH, Woolhandler S, Himmelstein DU. Giving office-based physicians electronic access to patients'

- prior imaging and lab results did not deter ordering of tests. *Health Aff* 2012; **31**: 488–96.
- 83 Bhangu A, Soreide K, DiSaverio S, Assarsson JH, Drake FT. Acute appendicitis: modern understanding of pathogenesis, diagnosis, and management. *Lancet* 2015; **386**: 1278–87.
 - 84 van der Meer MG, Backus BE, van der Graaf Y *et al*. The diagnostic value of clinical symptoms in women and men presenting with chest pain at the emergency department, a prospective cohort study. *PLoS ONE* 2015; **10**: e0116431. doi:10.1371/journal.pone.0116431.
 - 85 Epner PL, Gans JE, Graber ML. When diagnostic testing leads to harm: a new outcomes approach for laboratory medicine. *BMJ Qual Saf* 2013; **22**: ii6–10.
 - 86 Lo JA, Fisher DE. The melanoma revolution: from UV carcinogenesis to a new era in therapeutics. *Science* 2014; **346**: 945–9.
 - 87 Soda M, Choi YL, Enomoto M *et al*. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature* 2007; **448**: 561–6.
 - 88 Sparano JA, Gray RJ, Makower DF *et al*. Prospective validation of a 21-gene expression assay in breast cancer. *NEJM* 2015; **373**: 2004–15.
 - 89 Brenner S. Sequences and consequences. *Proc Trans R Soc B* 2010; **365**: 201–12.
 - 90 Tyler AL, Asselbergs FW, Williams SM, Moore JH. Shadows of complexity: what biological networks reveal about epistasis and pleiotropy. *BioEssays* 2009; **31**: 220–7.
 - 91 Podgornaia AI, Laub MT. Pervasive degeneracy and epistasis in a protein-protein interface. *Science* 2015; **347**: 673–7.
 - 92 Roberts NJ, Vogelstein JT, Parmigiani G, Kinzler KW, Vogelstein B, Velculescu VE. The predictive capacity of personal genome sequencing. *Sci Transl Med* 2012; **4**: 133ra58.
 - 93 Weiss KW, Terwilliger JD. How many diseases does it take to map a gene with SNPs? *Nat Genet* 2000; **26**: 151–7.
 - 94 Visscher PM, Hill WG, Wray NR. Heritability in the genomics Era – concepts and misconceptions. *Nat Rev Genet* 2008; **9**: 255–65.
 - 95 Manolio TA, Collins FS, Cox NJ *et al*. Finding the missing heritability of complex diseases. *Nature* 2009; **461**: 747–53.
 - 96 Joyner MJ, Prendergast FG. Chasing mendel: five questions for personalized medicine. *J Physiol* 2014; **592**: 2381–8.
 - 97 Hayden EC. Life is complicated. *Nature* 2010; **414**: 664–7.
 - 98 Elsasser WM. *The Physical Foundation of Biology*. New York: Pergamon, 1958.
 - 99 Brugts JJ, Yetgin T, Hoeks SE *et al*. The benefits of statins in people without established cardiovascular disease but with cardiovascular risk factors: meta-analysis of randomised controlled trials. *BMJ* 2009; **338**: b2376.
 - 100 Pursnani A, Massaro JM, Dagostino RB, O'Donnell CJ, Hoffman U. Guideline-based statin eligibility, coronary artery calcification, and cardiovascular events. *JAMA* 2015; **314**: 134–41.
 - 101 Gaziano JM, Gaziano TA. Simplifying the approach to the management of dyslipidemia. *JAMA* 2009; **302**: 2148–9.
 - 102 Brennan P, Croft P. Interpreting the results of observational research: chance is not such a fine thing. *BMJ* 1994; **309**: 727–30.
 - 103 Austin PC, Mamdani MM, Juurlink DN, Alter DA, Tu JV. Missed opportunities in the secondary prevention of myocardial infarction: an assessment of the effects of statin under-prescribing on mortality. *Am Heart J* 2006; **151**: 949–55.
 - 104 Emami-Riedmaier A, Schaeffeler E, Nies ET, Möirike K, Schwab M. Stratified medicine for the use of antidiabetic medication in treatment of type ii diabetes and cancer: where do we go from here? *J Int Med* 2015; **277**: 235–47.
 - 105 Hand D, Mannila H, Smyth P. *Principles of Data Mining*. Cambridge: MIT, 2001.
 - 106 Medical Academia Is Failing Patients and Clinicians By. Neglecting basic observational clinical research. *BMJ* 2006; **332**: 863–4.
 - 107 Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? *BMJ* 2014; **348**: g3725.
 - 108 Califf RM, Platt R. Embedding cardiovascular research in practice. *JAMA* 2013; **310**: 2037–8.
 - 109 Katzen IL, Rudick RA. Time to integrate clinical and research informatics. *Sci Transl Med* 2012; **4**: 162 fs41.
 - 110 Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care. *Nature Revs Genetics* 2013; **14**: 75.
 - 111 Liao KP, Cai T, Savova GK *et al*. Development of phenotype algorithms using electronic medical records and incorporating natural language processing. *BMJ* 2015; **350**: h1885.
 - 112 Hill RG, Sears LM, Melanson SW. 4000 Clicks: a productivity analysis of electronic medical records in a community hospital ED. *Am J Emer Med* 2013; **31**: 1591–4.
 - 113 Adler-Milstein J, Huckman RS. The impact of electronic health record use on physician productivity. *Am J Managed Care* 2013; **19**: SP345–52.
 - 114 Turchin A, Goldberg SI, Breydo E, Shubina M, Einbinder JS. Copy/paste documentation of lifestyle counseling and glycemic control in patients with diabetes: true to form? *Arch Int Med* 2011; **171**: 1393–4.
 - 115 Abelson R, Creswell J, Palmer GJ. Medicare bills rise as records turn electronic. *NY Times* 2012; nytimes.com/2012/09/22/business/medicare-billing.
 - 116 Office of the Inspector General. Early Assessment Finds That CMS Faces Obstacles in Overseeing the Medicare EHR Incentive Program. <https://oig.hhs.gov/oei/reports/oei-05-11-00250.asp>.
 - 117 Berner ES, Kasiraman RK, Yu F, Ray MN, Houston TK. Data quality in the outpatient setting: impact on clinical decision support systems. *AMIA Annu Symp Proc* 2005; 41–5.
 - 118 Kazmi Z. Effects of exam room EHR use on doctor–patient communication: a systematic literature review. *Inform Prim Care* 2013; **21**: 30–9.

Correspondence: D. Zakim, Unit for Bioentrepreneurship (UBE), Medical Management Centre at the Department of Learning Informatics, Management and Ethics (LIME), Karolinska Institutet, Stockholm, Sweden.
(e-mail: david.zakim@ki.se). ■