ABSTRACT
The widespread adoption of commercial electronic health records (EHRs) presents a significant challenge to the field of informatics. In their current form, EHRs function as a walled garden and prevent the integration of outside tools and services. This impedes the widespread adoption and diffusion of research interventions into the clinic. In most institutions, EHRs are supported by clinical operations staff who are largely separate from their informatics counterparts. This relationship needs to change. Research informatics and clinical operations need to work more closely on the implementation and configuration of EHRs to ensure that they are used to collect high-quality data for research and improvement at the point of care. At the same time, the informatics community needs to lobby commercial EHR vendors to open their systems and design new architectures that allow for the integration of external applications and services.

A CHANGING LANDSCAPE
The field of informatics faces two fundamental, but related challenges. First, the widespread adoption of commercial EHR systems presents a barrier to researchers and investigators seeking to deploy or integrate new tools and interventions into clinical practice. Second, despite their challenges, EHRs will remain the primary method of collecting/generating clinical data—yet there is little to no input from the informatics community into their configuration, implementation, and support. For the field of informatics to remain relevant, successful tools and interventions that are developed for research need to be put into general clinical practice and integrated into the clinical workflow. This is one of the primary goals of the Patient-Centered Outcomes Research Institute (PCORI), a tool into the EHR, however, proved to be almost impossible. There was no way to recreate the dynamic, self-reported survey tool or import the results, the EHR was not capable of computing SDs for the laboratory values, and it lacked functionality for integrating the adherence data. In the end, most of the decision support and risk stratification was jettisoned, leaving only the calculation of the adherence score, which had to be manually entered into the EHR.

As another example, my development team is halfway through a federal grant where we are working in collaboration with ImproveCareNow, a network of 36 centers focused on pediatric inflammatory bowel disease (IBD) that is aiming to build an ‘EHR-linked’ registry. Data are collected in discrete fields within the EHR during the clinic visit, extracted, and then uploaded into the registry. They could then be used for research, quality improvement, and clinical purposes such as previsit planning, population management, and patient engagement. An EHR-linked registry would achieve the long-cherished clinician goal of ‘data in once,’ that is, data initially collected for clinical care and then reused for other purposes. The ImproveCareNow network is a particularly good candidate for such a project as their registry data are the same elements that should be collected as part of model IBD care.
When we started the project, we approached the various EHR vendors represented in the network to ask them about creating and supporting a form that could be used by all of their customers. Aside from Epic, which quickly agreed to participate, the other EHR vendors have been very slow to respond. As a result, all of the non-Epic centers are likely to have to build and support their own data collection forms and create their own custom extracts, which poses a barrier to adoption. Since few institutions have EHR staff dedicated to handling research requests, this project must compete for priority against initiatives like the ICD-10 conversion and Phase I of Meaningful Use or the implementation of the EHR itself. Once we do have the attention of the EHR staff, we must ensure that the form is built so that it: (a) easily integrates into the clinical workflow (something frequently ignored by researchers); (b) can be completed quickly; (c) can, once completed, be pulled into a progress note or referral letter (also overlooked), and (d) has elements that are easily queried/extracted. If criteria a, b, and c are not met, then the clinicians will not use the form. If criterion d is unsatisfied, the enterprise reporting team will have trouble generating an extract. With EHRs, the easiest method of data collection is to typically create a text-based template where a clinician can fill in certain keywords. This leads most clinicians to believe they are capturing data discretely, but unless a significant amount of effort is undertaken to tag those keywords with a unique identifier, the result is stored as a blob of text. Other data collection methods may store data discretely, but do not integrate into the clinical workflow. The design of EHRs should be such that making the ‘right’ choice in terms of designing data collection forms is the easy choice.

Getting data into the health record is only half the challenge. After the forms are created, we face a second set of issues in transferring the data back out. In this project, we have taken the decidedly low-tech approach of having each center work with their enterprise reporting team to generate a report, which will then be uploaded to the registry. Two of the more sophisticated approaches to this problem, data transfer using the Health Level 7 (HL7) Continuity of Care Document (CCD) or the Integrating the Healthcare Enterprise (IHE) Retrieve Form for Data Capture (RFD) protocol, while appealing, are not suitable for our purposes. With the HL7 CCD, none of the condition-specific variables needed by ImproveCareNow are included in the specification. With RFD, an emerging standard that allows clinical trial forms to be embedded in the EHR and pre-populated with certain data, the data in the form are sent to a remote server. No data are retained in the EHR, however, which is a significant drawback when the data needed for research are also utilized for clinical care. In the end, far too much effort is spent working around the limitations of the EHR instead of developing better tools and interventions.

**WE NEED A DIFFERENT MODEL**

For EHRs to be truly useful and not just an electronic version of a paper chart, their design must change. It must be easier to create data collection forms and other interfaces that are easily integrated into the clinical workflow yet are tied to underlying data elements that are common across the vendor’s system. Customers should be able to author these forms, but there should be a mechanism that allows them to be vendor supported through system extensions and upgrades. At the same time, it needs to be easier to extract information from EHRs through common application programming interfaces (APIs). We need alternatives to the current approaches of a database query or an HL7 feed. In addition to making data sharing easier, this would also allow innovation in areas like safety monitoring and decision support, which require real-time access to production EHR data. While it may take some time, the winning EHR design of the future will be one that adopts a robust data model with a loosely coupled service-oriented architecture. This would allow easier access to the underlying data and an interface for the use of external services such as quality reporting and decision support. While the ideal case would be to have all EHRs utilize a common reference model with a corresponding set of end-user APIs, that is an unlikely prospect. A more realistic and obtainable solution would be to have all of the EHR vendors publish their APIs (and create them if they do not already exist). Having standardization across vendor APIs for a given functional area (demographics or medication orders, for instance) would be optimal, but one could certainly imagine the emergence of a secondary market that would provide such standardized overlays to achieve the same result.

Having a published set of APIs is the first step towards the establishment of an ecosystem of external services and applications. Given the current environment, it is unrealistic to expect a completely wide-open marketplace like that of Android. A more step-wise solution would be to establish a set of design patterns and an approval process for developers to deploy their tools (ie, the Apple approach). To some degree, this is the concept behind the Substitutable Medical Apps, reusable technologies (SMArt) project, although unlike the current version of SMArt, the ability to write data would be required. There are definite limitations to the Apple model, but even so, it would be a drastic improvement over what we have today. In this landscape, the vendor’s revenue stream would come from de-identified comparative effectiveness research or other analysis of their patient population. If vendors also adopted the Apple philosophy toward external applications, they could establish a secondary revenue stream through licensing fees. Without outside pressure from the marketplace, from patients, and from clinicians, however, vendors have no incentive or motivation to change. Perhaps AMIA should convene a working group to liaise between the informatics community, the general public, and commercial EHR vendors, providing a forum for a more open and transparent dialogue. Barré such a development, innovation in medical informatics will occur at the speed of the vendor’s upgrade cycle.

While the informatics community needs to push for change in the design of EHRs, at the same time it needs to become more involved in their configuration and implementation. For years, the relationship between clinical operations and research at most institutions ranged anywhere from non-existent to hostile, to, in the best cases, something close to benign neglect. The implementation of enterprise-wide EHRs has now caused these worlds to come crashing together. It is time for research and operations to become more integrated. The operational staff who are charged with implementing the EHR have the most control over the design and configuration of the final build. They are the ones who know the capabilities of the system and are the ones making recommendations on how to electronically capture the outputs of a clinical workflow. If we are to ever achieve the vision of a learning health system, where learning occurs with every patient encounter, we need to ensure that the clinical information systems are configured to allow this learning. Otherwise, we should not be surprised when EHR data turn out to be less than optimal for quality improvement and research. We must leverage the vast numbers of staff who are tasked with maintaining the institutional...
EHR. Expectations must be set that their role is not only to support clinical care, but also quality improvement and research. In a learning health system, all three occur in every single encounter and cannot be separated. The clinical informatics community has much to contribute here, particularly in the areas of clinical decision support, care process redesign, and workflow analysis/optimization. In the same way that other faculty and staff run laboratories and/or see patients, we should embrace (and reward) this service role. By leveraging our expertise in clinical terminology representation, structured data capture, interface design, etc, we can drive towards a more successful implementation and utilization of existing commercial EHRs. This type of collaboration is messy and requires a tremendous amount of effort at times, but given a world where government funding appears increasingly threatened, is likely to be the key to survival.

Funding This manuscript references work that was funded in part by R01HS20024 from the Agency of Healthcare Research and Quality and an internal grant award from the Cincinnati Children’s Hospital Medical Center.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

Correction notice This article has been corrected since it was published Online First. It has been made unlocked.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/3.0/