Modern Information Technology for Cancer Research: What's in IT for Me? An Overview of Technologies and Approaches

Janko Ahlbrandt Martin Lablans Katrin Glocker Sophia Stahl-Toyota
Klaus Maier-Hein Lena Maier-Hein Frank Ückert

German Cancer Research Center, Heidelberg, Germany

Keywords

Research informatics · Cancer research · Secondary use ·
Collaboration · Machine learning · Mobile health

Abstract

Information technology (IT) can enhance or change many scenarios in cancer research for the better. In this paper, we introduce several examples, starting with clinical data reuse and collaboration including data sharing in research networks. Key challenges are semantic interoperability and data access (including data privacy). We deal with gathering and analyzing genomic information, where cloud computing, uncertainties and reproducibility challenge researchers. Also, new sources for additional phenotypical data are shown in patient-reported outcome and machine learning in imaging. Last, we focus on therapy assistance, introducing tools used in molecular tumor boards and techniques for computer-assisted surgery. We discuss the need for metadata to aggregate and analyze data sets reliably. We conclude with an outlook towards a learning health care system in oncology, which connects bench and bedside by employing modern IT solutions.

© 2018 S. Karger AG, Basel

Introduction

Fighting cancer is the challenge of a lifetime for patients, their physicians and researchers in oncology. Hence, accepting all help available seems only logical. And where would one expect help in an era of digital disruption [1] if not in information technology (IT)? However, despite big promises to improve patient care and reduce errors or work load [2], using data and learnings from basic and clinical cancer research remains a challenge [3]. Even beyond scientific content, using IT tools today requires technical expertise and infrastructure, which is complex and costly to set up and operate. In this article, we give you a few impressions of scenarios employing technologies to support and improve cancer research or therapy. Following a prototypical process of research, we take you from solutions to explore a field by data, generate hypotheses and identify or share existing data within research networks. We advance to gathering and analyzing genomic data, opening up new data sources with patient engagement or machine learning in imaging, touch on techniques for analytics and visualization and end with support for experimental or augmented therapies. We believe such a journey can neither be ex-

haustive nor representative, as there is no registry for IT tools or specific key words to search for. However, we hope to give you – the oncologist or scientist in oncology – a few ideas on how to get in the driver's seat on the road to better outcomes.

Both oncologic treatment and research are for the most part still considered a data-poor environment [4], as most data is not systematically (and electronically) gathered, structured or integrated into IT systems. There are small areas that involve large data sets (e.g., omics data or radiology data), but data on processes, results or external knowledge still requires a lot of manual work or is not recorded at all.

Other sciences (e.g., meteorology and astrophysics) already have infrastructures in place to collect and share data because they invested in them [4]. To transform biomedical research – and cancer research in particular – into a data-rich environment, it will take a strong commitment to empower people and infrastructure. Many of the necessary building blocks already exist or are currently being researched – as is outlined below. However, there are many other areas that need more attention, especially by physicians driving innovation for the sake of their patients. One big area to focus on should be electronic health records (EHRs). They need to go beyond representations possible on paper and provide structured information in order to become useful for research [5]. A few good examples are already out there, as some companies develop systems for oncologists to use for both care and research [6].

Existing IT solutions will not cure patients by themselves; they need to be employed by caretakers in all medical professions. Many advances in medical IT research are yet to be introduced into routine care. Some apprehensions about investing in health IT are that more reliance on IT could introduce error [7], divert attention to other patients [6] or simply interfere with workflows [8]. However, if implemented wisely, it can open a window of opportunity for faster research, more insights and therapy options and ultimately a better outcome for patients.

Impressions of IT Usage Scenarios in Cancer Research

Data Extraction and Representation for Scientific Use

Starting our journey towards new insights, we want to explore the field of oncology to identify areas to improve cancer therapy or form hypotheses for pathogenesis. Data on current diagnostics, therapies and their outcomes

could provide a great start. Information on patients collected during diagnosis or treatment is usually stored in an EHR. Even though EHRs are not designed for research, they represent an unprecedented source of patient information for research if properly augmented [4]. Naturally, it has long been a goal to make this data available for research and many attempts have already been made in many projects [9–11]. Unfortunately, poor data quality and quantity remain a challenge [12, 13]. It is possible to extract clinical information from EHRs using natural language processing [14]. However, as of this writing, extracted information can be produced at scale, but is rarely as useful as structured information [15]. Hence, improving documentation of structured data and thereby improving data quality remains the best option to gain usable data for research. However, quantity can still remain a challenge. But it can be overcome by sharing data, for example, within a research network.

Collaboration in Research Networks

Research sites and institutions will always have limited resources like staff, facilities, study participants or samples. Therefore, it seems logical to cooperate with other researchers to form a research network. Within these networks, knowledge and samples can be exchanged, but data exchange seems even more crucial. Whether to increase the sample size in a study or to compare different studies performed, more data promises to improve results dramatically. Building such networks is possible, as has been proven by the German Cancer Consortium [16]. However, integrating data documented at independent institutions is a massive undertaking, especially if it was originally documented for treatment purposes [17]. Since the data to be shared is found in various documentation systems, a multitude of technical interfaces have to be developed [18] (syntactic interoperability). And as staff at different sites have slightly different conventions to document findings, mapping the various meanings to a common data set is far from trivial (semantic interoperability). Data privacy is also an issue, as the data relates directly to cancer patients. Well-established IT solutions for pseudonymization are freely available [19] but have to be integrated into daily processes by each institution. For example, obtaining the informed patient consent required for most forms of data processing [20] (EU-GDPR) occupies the physician's already scarce time. A decentralized data storage concept employing centralized metadata and well-defined processes for data access can help to comply with regulation and still yield respectable results.

Cloud-Based Collaboration

Though controversial from a data privacy standpoint [21], the advantages of cloud computing are enticing: virtually unlimited resources for storing data, analyzing and visualizing it on demand plus instant access for everyone authorized [22]. The Seven Bridges Cancer Genomics Cloud is an example for such a cloud to analyze genomic databases without the need for a local infrastructure to hold the large data sets [23]. By providing access to these data sets from different sources, the cloud can enable machine learning or finding similar cases, as they (generally) rely on a large number of samples. IBM Watson, for instance, could identify drug candidates by sifting through massive amounts of data in pilot studies [24]. Beyond storing data, cloud computing can also execute algorithms supplied by researchers directly on the data stored [22, 23]. This enables other researchers to use the algorithms and gain insights from their output without the need for deep technical knowledge [25]. Lastly, by strictly versioning the data sets, the algorithms used and the infrastructure, the results of any given combination are reproducible [25]. Looking back on our scientific journey, we have now covered sharing and access to data. Next up is analysis – focusing on genomic data.

Analyzing Genomic Information

Genomic diagnostics promises to change cancer research and ultimately cancer therapy for good. So, it is likely that most cancer research will incorporate it in the future. Even today, therapy for cancer patients is guided by findings in their tumor genome [26]. However, the quantity and quality of proven causal relationships between genome changes and clinical implications are still low [26] compared to centuries of “conventional” medical knowledge. Phenotype information is needed to improve research on genome-related diagnostics and to create a learning healthcare system [27]. Unfortunately, obtaining both genomic and phenotype data from the same patients can be very difficult. Additionally, phenotypic data is often of low quality (i.e., incomplete or imprecise) or is not semantically comparable (i.e., the same condition is recorded in different ways). Ultimately, for patients to benefit from genetic diagnostics, a physician has to correctly interpret the results and draw conclusions [26]. This is by no means a small task and, besides costs, probably the biggest obstacle towards widespread adoption. Hence, we need more certainty in predictions based on genetic data and coupling it with large amounts of other patient data could help to fill that void.

We set out our journey to gather information on the outcome of cancer therapy, which is ultimately the condition of patients. Involving them in their care and opening up new data sources for research is next.

Patient Engagement and Reported Data

It takes innovative approaches in cancer therapy a long time to be accessible to physicians and patients. One underlying issue in cancer treatment research is the lack of trials, more specifically the lack of participation therein. Patient participation in trials is low compared to the number treated (in the US 2–3% of adults) [6]. Besides other concerns, like fear of placebo-treatment for instance, the effort for participants seems to be a major barrier. Well-designed informatics approaches could help with that. A good example can be drawn from cardiology: Data gathered from wrist-worn devices or from patient reports through their smartphones was included in a study in a large scale [28, 29]. Increasing patients' engagement in their treatment can improve their experience and outcome [30]. It can also yield valuable data for cancer research.

mHealth – health information and services provided through mobile devices – can increase patient engagement; however, there are many challenges, either technical, legal or human to overcome [31]. Data on the quality of life for patients (as measured by standardized methods like the 36-Item Short Form Survey [32] or EuroQol-5 Dimensions [33]) – especially those undergoing cancer treatment – could form a valuable basis to enhance treatment decisions [34]. First, results promise psycho-oncological benefits for patients with cancer [35]. Beyond informing patients, mobile devices also allow them to share information on their physical or psychological status and needs. Evaluating the outcome or current condition of patients beyond mortality has recently received more attention in oncology, especially by health payors or insurers [34]. This attention is highly warranted, since integrating patient-reported symptoms into clinical care could increase survival for patients with metastatic cancer [36]. Enabling patients to submit their information to their physicians or caretakers could also make it available for research. Getting information straight from the source like this can highly increase the quantity of data-points. For instance, a detailed pain history (with many data-points) can indicate an insufficient pain medication in comparison to a single value at a scheduled appointment. In the United Kingdom, patients report on their results and status to assess the quality of certain surgeries performed [34]. Altogether, patient-reported information

could be very valuable for cancer research. But there are more sources we can tap, as we continue to derive data from images.

Machine Learning to Detect Cancer in Images

Image-based diagnostics currently focusses on machine learning based on deep convolutional neural networks (CNN) and radiomics. Radiomics [37] aims to extract quantitative, reproducible, structured, machine readable information out of medical images. Combining the 2 can have a significant impact: One example improves mammography screening by significantly reducing the false positives findings and therefore avoiding unnecessary invasive biopsies [38]. The system can automatically detect breast lesions on diffusion-weighted MR images. Another example is the improved stratification of Glioblastoma regarding overall survival and progression-free survival by adding radiomic subtyping to molecular, clinical or standard imaging parameters [39].

Other applications aim to assist in identifying image pattern in microscopy to reveal hints for cancer. A group at Google created an augmented reality application that uses machine learning to highlight suspicious areas in samples [40]. They trained a model to extract these areas based on sample areas previously identified as cancer. This example shows the huge potential that exists in combining computers, which are good at recognizing patterns consistently, and humans, who can check results for plausibility and reason on them. Ultimately, Radiomics opens up the treasures of data buried in medical images to research and analysis for both humans and machine learning algorithms.

Getting a Grip on Information Complexity and Load

So far, our journey through IT in cancer research took us to analyzing data and augmenting it with more data from innovative sources. Making sense of vast amounts of multiform data – a process described as analytical reasoning [41] – is now required to derive new insights. Out of all the findings in clinical information – patient-reported results, genetic variants, imaging data – which ones are significant? And how do they relate to studies, case reports, medication trials or cellular pathways? Research and good examples hint at what is possible with visualization [41–43]; however, there is no report of widespread use of these technologies in real-world environments [43]. Especially within a research context, visualizations are a chance to clean data sets from error-laden data or detect patterns more easily [44]. Also, by means of visualization, one can explore a data set to un-

derstand unexpected trends or hints, even without a sophisticated analysis [45]. In a research setting, this is very valuable for hypothesis generation or preliminary analysis (e.g., for feasibility studies). Remarkably, domain scientists (like cancer researchers) have trust in – and will hence likely benefit quickly from – newly introduced visual analytics tools when they are intuitive, transparent and controllable [46]. Their level of trust in visual tools is even higher than that in manual analysis methods when performing complex tasks [46]. Thus, visualization has benefits beyond efficiency gains and could provide researchers with better insights. As these should quickly find their way into therapy, we propose to cover therapy assistance next.

Therapy Assistance with Omics and External Knowledge

Personalized diagnostics based on genomic data is in high demand, since changes in genomic material can have huge effects [47–49]. Hence, Molecular Tumor Boards are established for biologists, pathologists and oncologists to collaborate in finding targeted therapies for patients, who do not benefit from standard treatment [50]. For the Molecular Tumor Boards at the National Center for Tumor Diseases [51] in Germany bioinformatics pipelines analyze patient and cancer cell transcriptome and exome, search for modifications and filter for those known to be susceptible for medication [52]. Using this information, physicians gather known publications, past or present trials, known gene relations and further knowledge on specific variations or mutated genes. Most tools are freely available web applications and do not require installation and are therefore easily accessible:

cBioPortal [53] provides an overview of gene variation and their domains. The public version accumulates data from currently 248 studies. Like with similar tools [54], one can search for one or more specific genes to obtain information on studies, gene relationships or clinical relevance. CIViC (Clinical Interpretations of Variants in Cancer: <https://civicdb.org/home>) also focusses on clinical relevance. One can query for gene-wide or specific variants and receive precise information on variant types and drugs in various phases, from in-vitro studies to results of phase 3 clinical trials.

Information about completed studies is often found in literature databases, such as PubMed. Queries can be optimized using MeSH (Medical Subject Heading) Thesaurus, which is used to index all articles in PubMed [55]. Embase (<https://www.embase.com>) by Elsevier is a pay-to-use database and has the advantage of intuitive search

options like PICO, allowing the option of filtering by population, intervention, comparison and outcome.

As the number of studies focusing on molecular markers is increasing steadily, ClinicalTrials.gov (<https://clinicaltrials.gov/>) can list studies to enroll the patient in or learn from their – sometimes even preliminary – findings. This manual task of matching a patient to trials (finished or ongoing) is complex [56] but could be automated in the future [57].

Computer-Assistance for Cancer Surgery

Surgical data science aims to use all available data of the care process and integrate it to assist physicians and caregivers [58]. As machine learning is applied to solve specific problems, the shortage of curated data sets becomes a bottleneck. In minimally invasive surgery, for example, it is extremely challenging to capture all the variations (e.g., all medical devices applied in an intervention) that may occur in clinical practice. Furthermore, algorithms trained on data of one specific domain (e.g., specific hardware for a specific intervention) typically cannot be generalized well for new domains. Approaches proposed to overcome these challenges include the following:

1. Crowdsourcing: To spare resources of (medical) experts, annotation tasks may be outsourced to masses of untrained workers from an online community [59, 60]. The workers' clickstreams (e.g., pattern of cursor movement, number of clicks/zooms) can be automatically analyzed [61] to estimate the quality of annotation. Also, hybrid human-crowd annotations concepts have been proposed for large-scale medical image annotation at low cost [62].

2. Self-supervised learning: While annotated medical data is rare, unlabeled medical data is much more easily accessible. Based on this observation, Ross et al. [63] investigated the hypothesis that masses of unlabeled video data can be used to learn a representation of the target domain that boosts the performance of state-of-the-art machine-learning algorithms. In other words, they showed that CNNs can even train themselves to recognize structures within video material, given that they were previously exposed to example images of the same domain.

Discussion and Outlook

At the end of our journey, many, if not all, of the presented scenarios, specialized systems or platforms provide information or answers for individual questions. Unfortunately, cancer research or treatment yields com-

plicated questions – and answers – at high speed. Maintaining up-to-date information and processing new information are daunting tasks. The sheer volume of information to query and process is at some point destined to overwhelm a human's capabilities – no matter how smart or educated. Hence, the challenge is not just in making information available but also in the capability of processing a vast amount of information and filter for actionable insights.

Combining some of the techniques shown, like data extraction, visualization or inclusion of genomic data could speed up hypothesis generation for research. However, in order to yield valid results, we need high data quality and good metadata – data describing data. Metadata can ensure that data sets – or rather their content – are valid and comparable. For instance, by enforcing the same vocabulary to describe certain conditions or provide a way to convert them.

Once supplied with good data and metadata, systems for presenting it have to be crafted in order to reduce the complexity. Since the types of questions and hence analyses differ between researchers, they each require their own system. Unfortunately, there is no “plug-and-play” system for research data available. Some portals combine data sets or offer different workflows, but they are not truly customizable (yet).

Still, you can take the next step, if you are willing to learn (and invest some time). Engage in improving EHRs and other documentation systems, so that entered data can be reused for scientific purposes! We believe that open data sharing and improving data quality can yield exceptional results and are first steps towards a “learning health care system” [3, 64]. But it will require commitment, investment and the willingness to cooperate of many researchers involved. As long as control over a specific data set yields a competitive advantage and scientific partners are not trusted, the system at large will not benefit as much as it should.

Despite all the obstacles outlined, cancer research is en route to a data rich environment, which could enable a much faster turn-around time. In common life science studies, data is collected to test one specific hypothesis and rarely reused. This process is very labor and time intensive. Preexisting data-pools and well-curated metadata could enable testing a hypothesis in about a day's time – or even less. Machine-learning approaches could speed this up even more or broaden the scope of analyses. We believe data-rich cancer research will empower researchers to spend their time more effectively. Basic research will always be necessary, but it could be guided

more effectively. Data- and knowledge-driven clinical cancer research could help to bring innovation from the labs to the patients a lot faster.

Just like a person is much more than the sum of her parts, treating a patient with cancer needs a holistic view. Considerable progress has been made to understand individual processes and tools have been crafted to unlock previously unknown secrets. But putting the pieces together requires the oncologist, who sees her patient as a whole and is determined to find the best possible therapy

by incorporating all that modern science has to offer. Modern IT tools can be a way to create and leverage scientific insights and bring them to the bedside, just as it can supply researchers with valuable data about real patients.

Disclosure Statement

The authors declare that they have no conflicts of interest to disclose.

References

- Gilbert RJ. E-books: A tale of digital disruption. *J Econ Perspect*. 2015;29(3):165–84.
- Lambin P, van Stiphout RG, Starmans MH, Rios-Velazquez E, Nalbantov G, Aerts HJ, et al. Predicting outcomes in radiation oncology – multifactorial decision support systems. *Nat Rev Clin Oncol*. 2013 Jan;10(1):27–40.
- Kibbe W. Cancer Clinical Research. *Oncology Informatics*. Elsevier; 2016. pp. 41–53.
- Riley WT. A New Era of Clinical Research Methods in a Data-Rich Environment. *Oncology Informatics*. Elsevier; 2016. pp. 343–55.
- Miriovsky BJ, Shulman LN, Abernethy AP. Importance of health information technology, electronic health records, and continuously aggregating data to comparative effectiveness research and learning health care. *J Clin Oncol*. 2012 Dec;30(34):4243–8.
- Hesse BW, Ahern D, Beckjord E. Oncology Informatics: Using Health Information Technology to Improve Processes and Outcomes in Cancer. San Diego, UNITED STATES, Elsevier Science & Technology, 2016. Available from: <http://ebookcentral.proquest.com/lib/dkzf/detail.action?docID=4454380>.
- Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, et al. Role of computerized physician order entry systems in facilitating medication errors. *JAMA*. 2005 Mar;293(10):1197–203.
- Abdiwahab E, Taplin SH, Coronado G, Dacus H, Leyboldt M, Skinner C. Early Detection in the Age of Information Technology. *Oncology Informatics*. Elsevier; 2016. pp. 123–43.
- Rea S, Pathak J, Savova G, Oniki TA, Westberg L, Beebe CE, et al. Building a robust, scalable and standards-driven infrastructure for secondary use of EHR data: the SHARPn project. *J Biomed Inform*. 2012 Aug;45(4):763–71.
- Safran C, Bloomrosen M, Hammond WE, Labkoff S, Markel-Fox S, Tang PC, et al.; Expert Panel. Toward a national framework for the secondary use of health data: an American Medical Informatics Association White Paper. *J Am Med Inform Assoc*. 2007 Jan-Feb;14(1):1–9.
- Linder JA, Haas JS, Iyer A, Labuzetta MA, Ibara M, Celeste M, et al. Secondary use of electronic health record data: spontaneous triggered adverse drug event reporting. *Pharmacoepidemiol Drug Saf*. 2010 Dec;19(12):1211–5.
- Botsis T, Hartvigsen G, Chen F, Weng C. Secondary Use of EHR: Data Quality Issues and Informatics Opportunities. *AMIA Jt Summits Transl Sci Proc*. 2010 Mar;2010:1–5.
- Feder SL. Data Quality in Electronic Health Records Research: Quality Domains and Assessment Methods. *West J Nurs Res*. 2018 May;40(5):753–66.
- Savova GK, Tseytlin E, Finan S, Castine M, Miller T, Medvedeva O, et al. DeepPhe: A Natural Language Processing System for Extracting Cancer Phenotypes from Clinical Records. *Cancer Res*. 2017 Nov;77(21):e115–8.
- Ford E, Carroll JA, Smith HE, Scott D, Cassell JA. Extracting information from the text of electronic medical records to improve case detection: a systematic review. *J Am Med Inform Assoc*. 2016 Sep;23(5):1007–15.
- Lablans M, Schmidt EE, Ückert F. An Architecture for Translational Cancer Research Exemplified by the German Cancer Consortium. *JCO Clin. Cancer Inform*. 2018:1–8.
- Kahn MG, Callahan TJ, Barnard J, Bauck AE, Brown J, Davidson BN, et al. A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data. *EGEMS (Wash DC)*. 2016 Sep;4(1):1244.
- Lablans M, Kadioglu D, Muscholl M, Ückert F. Exploiting Distributed, Heterogeneous and Sensitive Data Stocks while Maintaining the Owner's Data Sovereignty. *Methods Inf Med*. 2015;54(4):346–52.
- Lablans M, Borg A, Ückert F. A RESTful interface to pseudonymization services in modern web applications. *BMC Med Inform Decis Mak*. 2015 Feb;15(1):2.
- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons
- with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). 2016, [cited 2018 May 13]. Available from: <http://data.europa.eu/eli/reg/2016/679/oj/eng>.
- Chen D, Zhao H. Data Security and Privacy Protection Issues in Cloud Computing. In: 2012 International Conference on Computer Science and Electronics Engineering. 2012. pp. 647–51.
- Stein LD. The case for cloud computing in genome informatics. *Genome Biol*. 2010;11(5):207.
- Lau JW, Lehnert E, Sethi A, Malhotra R, Kaushik G, Onder Z, et al.; Seven Bridges CGC Team. The Cancer Genomics Cloud: Collaborative, Reproducible, and Democratized – A New Paradigm in Large-Scale Computational Research. *Cancer Res*. 2017 Nov;77(21):e3–6.
- Chen Y, Elenee Argentinis JD, Weber G. IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research. *Clin Ther*. 2016 Apr;38(4):688–701.
- Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res*. 2016 Jul;44(W1):W3–10.
- Meric-Bernstam F, Johnson A, Holla V, Bailey AM, Brusco L, Chen K, et al. A decision support framework for genomically informed investigational cancer therapy. *J Natl Cancer Inst*. 2015 Apr;107(7):djv098.
- Rehm HL. Evolving health care through personal genomics. *Nat Rev Genet*. 2017 Apr;18(4):259–67.
- McConnell MV, Shcherbina A, Pavlovic A, Homburger JR, Goldfeder RL, Waggot D, et al. Feasibility of Obtaining Measures of Lifestyle From a Smartphone App: The MyHeart Counts Cardiovascular Health Study. *JAMA Cardiol*. 2017 Jan;2(1):67–76.

- 29 Shcherbina A, Mattsson CM, Waggott D, Salisbury H, Christle JW, Hastie T, et al. Accuracy in Wrist-Worn, Sensor-Based Measurements of Heart Rate and Energy Expenditure in a Diverse Cohort. *J Pers Med*. 2017 May;7(2):3.
- 30 Krist AH, Nease DE, Kreps GL, Overholser L, McKenzie M. Engaging Patients in Primary and Specialty Care. *Oncology Informatics*. Elsevier; 2016. pp. 55–79.
- 31 Mohammadzadeh N, Safdari R, Rahimi A. Cancer care management through a mobile phone health approach: key considerations. *Asian Pac J Cancer Prev*. 2013;14(9):4961–4.
- 32 Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992 Jun;30(6):473–83.
- 33 Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. *Appl Health Econ Health Policy*. 2017 Apr;15(2):127–37.
- 34 Nixon A, Wild D, Muehlhausen W. Patient Reported Outcomes: an Overview. Turin, SEEd Medical Publishers, 2016 [cited 2018 Apr 30]. Available from: <http://public.eblib.com/choice/publicfullrecord.aspx?p=4712333>.
- 35 Ernst J, Broemer L. Digital unterstützte Modelle in der Patientenbeobachtung. *Forum (Genova)*. 2018:1–6.
- 36 Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment. *JAMA*. 2017 Jul;318(2):197–8.
- 37 Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016 Feb;278(2):563–77.
- 38 Jäger PF, Bickelhaupt S, Laun FB, Lederer W, Heidi D, Kuder TA, et al. Revealing Hidden Potentials of the q-Space Signal in Breast Cancer. In: Descoteaux M, Maier-Hein L, Franz A, Jannin P, Collins DL, Duchesne S, editors. MICCAI 2017. Springer International Publishing. Medical Image Computing and Computer Assisted Intervention. 2017. pp. 664–71.
- 39 Kickingereder P, Neuberger U, Bonekamp D, Piechotta PL, Götz M, Wick A, et al. Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma. *Neuro Oncol*. 2018 May;20(6):848–57.
- 40 An Augmented Reality Microscope for Cancer Detection. Res Blog [cited 2018 Apr 27]. Available from: <https://research.googleblog.com/2018/04/an-augmented-reality-microscope.html>.
- 41 Thomas JJ, Cook KA. A visual analytics agenda. *IEEE Comput Graph Appl*. 2006 Jan-Feb; 26(1):10–3.
- 42 Shneiderman B, Plaisant C, Hesse BW. Improving Healthcare with Interactive Visualization. *Computer*. 2013;46(5):58–66.
- 43 Khan A, Mukhtar H, Ahmad HF, Gondal MA, Ilyas QM. Improving Usability through Enhanced Visualization in Healthcare. In: 2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS). 2017, pp 39–44.
- 44 Onukwugha E, Plaisant C, Shneiderman B. Data Visualization Tools for Investigating Health Services Utilization Among Cancer Patients. *Oncology Informatics*. Elsevier; 2016. pp. 207–29.
- 45 Horowitz TS, Rensink RA. Extended Vision for Oncology. *Oncology Informatics*. Elsevier; 2016. pp. 287–303.
- 46 Dasgupta A, Lee JY, Wilson R, Lafrance RA, Cramer N, Cook K, et al. Familiarity Vs Trust: A Comparative Study of Domain Scientists' Trust in Visual Analytics and Conventional Analysis Methods. *IEEE Trans Vis Comput Graph*. 2017 Jan;23(1):271–80.
- 47 Dancey JE, Bedard PL, Onetto N, Hudson TJ. The genetic basis for cancer treatment decisions. *Cell*. 2012 Feb;148(3):409–20.
- 48 Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011 Mar; 144(5):646–74.
- 49 Wong KM, Hudson TJ, McPherson JD. Unraveling the genetics of cancer: genome sequencing and beyond. *Annu Rev Genomics Hum Genet*. 2011;12(1):407–30.
- 50 Schwaederle M, Parker BA, Schwab RB, Fanta PT, Boles SG, Daniels GA, et al. Molecular tumor board: the University of California-San Diego Moores Cancer Center experience. *Oncologist*. 2014 Jun;19(6):631–6.
- 51 Horak P, Klink B, Heining C, Gröschel S, Hutter B, Fröhlich M, et al. Precision oncology based on omics data: the NCT Heidelberg experience. *Int J Cancer*. 2017 Sep;141(5): 877–86.
- 52 Perera-Bel J, Hutter B, Heining C, Bleckmann A, Fröhlich M, Fröhling S, et al. From somatic variants towards precision oncology: evidence-driven reporting of treatment options in molecular tumor boards. *Genome Med*. 2018 Mar;10(1):20.
- 53 Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013 Apr;6(269):pl1.
- 54 Kurnit KC, Bailey AM, Zeng J, Johnson AM, Shufean MA, Brusco L, et al. “Personalized Cancer Therapy”: A Publicly Available Precision Oncology Resource. *Cancer Res*. 2017 Nov;77(21):e123–6.
- 55 Lipscomb CE. Medical Subject Headings (MeSH). *Bull Med Libr Assoc*. 2000 Jul;88(3): 265–6.
- 56 Patel C, Cimino J, Dolby J, Fokoue A, Kalyanpur A, Kershenbaum A, et al. Matching Patient Records to Clinical Trials Using Ontologies. *The Semantic Web*. Berlin, Heidelberg: Springer; 2007. pp. 816–29.
- 57 Wicks P, Vaughan TE, Massagli MP, Heywood J. Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. *Nat Biotechnol*. 2011 May;29(5):411–4.
- 58 Maier-Hein L, Vedula SS, Speidel S, Navab N, Kikinis R, Park A, et al. Surgical data science for next-generation interventions. *Nat Biomed Eng*. 2017;1(9):691–6.
- 59 Maier-Hein L, Mersmann S, Kondermann D, Bodenstedt S, Sanchez A, Stock C, et al. Can Masses of Non-Experts Train Highly Accurate Image Classifiers? In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014. Cham: Springer; 2014. pp. 438–45.
- 60 Maier-Hein L, Kondermann D, Roß T, Mersmann S, Heim E, Bodenstedt S, et al. Crowdtruth validation: a new paradigm for validating algorithms that rely on image correspondences. *Int J CARS*. 2015 Aug;10(8): 1201–12.
- 61 Heim E, Seitel A, Isensee F, Andrusis J, Stock C, Ross T, et al. Clickstream analysis for crowd-based object segmentation with confidence. *IEEE Trans Pattern Anal Mach Intell*. 2017 Nov;1–1.
- 62 Maier-Hein L, Ross T, Gröhl J, Glocker B, Bodenstedt S, Stock C, et al. Crowd-Algorithm Collaboration for Large-Scale Endoscopic Image Annotation with Confidence. In: Ourselin S, Joskowicz L, Sabuncu MR, Unal G, Wells W, editors. *MICCAI 2016*. Medical Image Computing and Computer-Assisted Intervention. Cham: Springer International Publishing; 2016. pp. 616–23.
- 63 Ross T, Zimmerer D, Vemuri A, Isensee F, Wiesenfarth M, Bodenstedt S, et al. Exploiting the potential of unlabeled endoscopic video data with self-supervised learning. *ArXiv171109726 Cs* 2017 [cited 2018 Apr 24]. Available from: <http://arxiv.org/abs/1711.09726>.
- 64 Schilsky RL, Miller RS. Creating a Learning Health Care System in Oncology. *Oncology Informatics*. Elsevier; 2016. pp. 3–21.