Published in Proceedings of the 3rd International Workshop on Semantic Web Meets Health Data Management (SWH 2020), Athens, Greese, 2 November 2020, which should be cited to refer to this work.

Decentralized Management of Patient Profiles and Trajectories through Semantic Web Agents*

Jean-Paul Calbimonte[1], Davide Calvaresi[1], and Michael Schumacher[1]

University of Applied Sciences and Arts Western Switzerland (HES-SO), Sierre, Switzerland
{firstname.lastname}@hevs.ch

Abstract. The usage of healthcare data for analytics and patient applications has increased in recent years opening a number of technical, ethical and scientific challenges. Among these, those related to the management of personal and sensitive health data have been addressed through decentralized solutions for patient data, often implemented and modelled using distributed agents and semantic technologies. In this paper[1] we present a technical summary of our previous works in this area, comprising efforts to: (i) use ontology models to represent patient trajectories, (ii) employ agent-based architectures to model and employ decentralized patient data exchanges, (iii) define agent cooperation and negotiation strategies for healthcare data interactions, (iv) adopt semantic data models for privacy-aware agents, and (v) implement multi-agent systems for real-time healthcare data processing.

1 Introduction

An increasing number of patients face the challenge of having to follow treatments at home, often having to cope with complications and issues, having to rely on healthcare professionals only when the situation becomes critical. Typical examples include people suffering from chronic diseases, or conditions following cancer survival [10]. Under these circumstances, physicians and healthcare personnel need to personalize the treatment and take into account individual characteristics of each patient, e.g., understanding the details of previous medical encounters, interventions, medication, symptoms, and co-morbidities. Digitization of health care records and the use of hybrid care technologies including monitoring and virtual health coaching have showed a great potential for improving the quality of support for these patients. In fact, nowadays health records include a considerable amount of information, which is often not fully exploited for the benefit of patients.

The emergence of methodologies to characterize patient trajectories [1] has opened a way for systematic analysis of patient contextual data, including healthcare events, conditions, co-morbidities, emotional and social indicators, self-reported outcomes, and observations during and after treatment [8, 15]. However,

[1] This paper is a *Technical Summary* of previous published works.

so far there is still a need for comprehensive techniques that allow exploiting trajectory data using Machine Learning (ML) and other related AI techniques. Although numerous previous works have focused on prediction, classification and automated learning on specific types/parts of a patient health record, the key challenge lies on the combination of the different layers that constitute a trajectory.

In this work, we advocate the use of intelligent agents as the foundational information management entity for patient trajectories, in conjunction with semantic models to support data exchange. We summarize our previous efforts in this context, which constitute a solid basis for semantic-aware and decentralized management of patient trajectory data. We argue that patients can delegate the management of personal trajectory data to dedicated agents, which in turn can automatically negotiate and cooperate with other agents, for instance to share and aggregate anonymized data, to grant access to agents of medical staff, or to allow ML processing and prediction. The scientific contributions summarized in this work essentially identify two key research axes: patient trajectory heterogeneity, and decentralization of data and processing. The first dimension refers not only to the high diversity of potential data sources (EHR, sensor data, imaging, self-reported outcomes), but also to the heterogeneity of modelling and representation of trajectories. Regarding decentralization, given the distributed nature of data sources and processing units, as well as the necessity for patients to have full control over their personal data, it is mandatory to provide technological means for autonomous data management.

In the remainder of this paper, we introduce the different challenges addressed in this context (Section 2), before briefly describing how each of our previous works addresses them. Hence, we discuss our contributions on ontology-based modelling of trajectory data in Section 3, decentralized personal data in Section 4, agent cooperation and negotiation in Section 5, privacy-aware agents in Section 7, and real-time and stream processing of trajectory data in Section 6. We provide a discussion on perspectives and future research lines in Section 8.

2 Challenges in Patient Trajectory Data Management

Among the numerous challenges regarding patient trajectory management and its exploitation we focus on the following fundamental aspects:

- *Trajectory modelling.* The difficulty of exchanging heterogeneous data is of particular importance in the healthcare domain, and more so in trajectories where information pieces may span from demographics to radiology images, or multi-parametric sensor observations. A coherent trajectory model can only be constructed if it relies on solid foundations, linked to existing standard vocabularies in the medical realm. Moreover, these models must be machine-understandable in order to allow autonomous entities (i.e., agents) to take decisions based on their contents.
- *Decentralization.* Patient data is bound to be stored, managed and processed in a distributed manner. The provenance of trajectory data typically includes

different hospitals, clinics and medical cabinets, which are not expected to respond to central governance. Moreover, given the current regulations on personal health data [14], patients may now have full control over where and how their data is stored and/or processed. It is imperative to empower patients and participants and allow them to take informed decisions regarding their own data, as well as results, clinical studies and processing outcomes based on them.

- *Cooperation.* A comprehensive approach to the exploitation of patient trajectory data can only be possible if multiple decentralized parties exchange responsibility and processing/analytics duties. Given that trajectory data may be managed by different entities, it becomes necessary to provide the mechanisms that guarantee effective interactions among them. These cooperation schemes include participatory data collection, data reuse, collaborative processing, etc.
- *Data Volume & Velocity.* It is essential to consider not only the large volumes of data related to patient trajectories, but also the speed at which they are produced and processed. We underline the importance of providing the capability of responding to real-time constraints, as well as to the streaming nature of certain trajectory data sources, e.g., those provided by sensing and IoT devices.
- *Data Privacy.* None of the above challenges can be fully addressed if the sensitive nature of patient data is not taken into consideration. Privacy preservation as well as trustful interactions are absolutely necessary for any trajectory-based system to be acceptable according to current regulations regarding data protection [14].

In the next sections we indicate how our previous works addressed these challenges, fundamentally through agent-based models and semantic technologies.

3 Ontology-based Trajectory Modelling

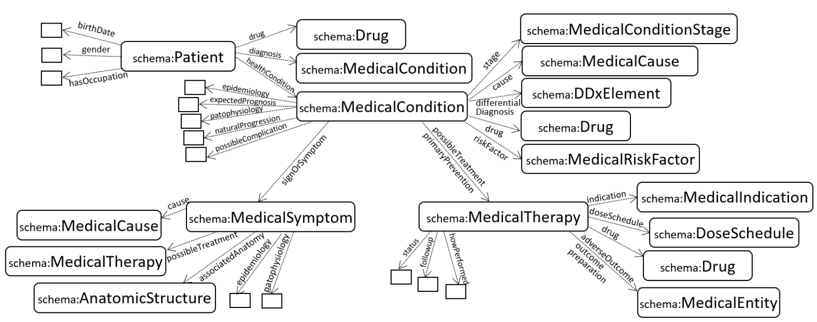

Fig. 1: Excerpt of relevant concepts for patient trajectories from schema.org [9].

In order to cope with the heterogeneity of patient trajectory data sources and types of data, we have proposed to represent them using semantic models linked to standard vocabularies in the medical, pharmaceutical and healthcare domain [4, 6]. Our approach regarding trajectory modeling follows an ontology-based methodology. A trajectory is conceived as a knowledge graph centered around the patient, and linking different entities that represent medical conditions, therapies, symptoms, drug prescription, diagnosis, monitoring observations, etc. We depict in Figure 1 a subset of such ontology model, based on the widely used schema.org [9] vocabulary. The implementation of this data model provides a solid basis for trajectory data exchange, with the possibility of mapping to standard formats such as HL7 FHIR. In parallel, our approach also emphasizes on the use on standard healthcare and bioinformatics ontologies. In our data model, we link schema.org coding attributes to standard terms in ontologies such as MeSH [11] or ICD-10 [12]. As a result, and as we will see later, a trajectory encoded as RDF following this model can be shared, processed and consumed by autonomous entities, according to established goals. Furthermore, explainable results can be represented using this semantic representation, providing the means for human-understandable trajectory analysis.

4 Decentralized Personal Data Management

One of the central contributions of our research on patient trajectories is the design of autonomous agents capable of managing personal health information. A main application for such agents is to provide behavior change support to patient with chronic conditions [3]. As depicted in Figure 2, we propose a model in which each user (patient) agent has exclusive access to personal data, providing with the ability to manage fine-grained access control and consent for any sharing or processing actions. Each agent is equipped with a *domain model* that describes as a knowledge graph the different factors that may affect behavior change. For example, this may include motivation or ability factors towards specific behavior change goals, e.g., adherence to a rehabilitation treatment. The agent also includes a set of stages, goals and milestones which can be incorporated to an internal *state machine*, as well as *computational persuasion* strategies that might be launched according to currently detected inputs and interactions. All these agent processes take into account the patient-specific information contained in the *user profile*, which is kept as part of the agent beliefs, and is typically composed of trajectory information modeled as described in Section 3.

Beyond the patient private sphere, the model also considers the following additional agents: (i) coordinator: it oversees the incorporation of new patient agents, as well as governing over stage changes and regulating interactions and disseminating domain model information that may affect groups of participants in a behavior change program; (ii) coach: it provides specialized guidance to a patient through his user agent, demanding access to certain personal data if it requires to be processed by a behavior agent; (iii) mediator: allows patient-to-patient interactions that may be useful in order to adopt peer-to-peer strategies for behavior change; (iv) behavior agent: in charge of providing trajectory and behavior pattern analysis through AI exploitation of anonymized data.

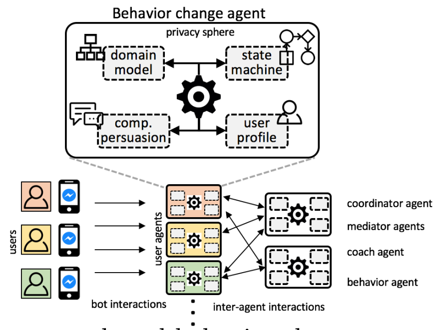

Fig. 2: A model for agent-based behavior change support applications [3].

5 Inter-agent Cooperation and Negotiation

Beyond modelling the agents themselves, it is critical to provide well-defined strategies for cooperation among patient (And trajectory) agents, so that they are able to interact as described in the previous section. We introduced in [6] the concept of τAgents, or patient trajectory management agents (Figure 3), which are characterized by a set of goals, beliefs, and behaviors; and include specialized knowledge graphs of patient trajectory data. τAgents may play different roles and accordingly adopt different (and even competing) goals and behaviors. For instance, a patient agent may focus on improving quality of life indicators, or to retain moderate physical activity over time. In contrast, a coaching agent may define its goals in terms of level of adherence of its assigned patients a certain therapy. The same logic applies to beliefs and behaviors, and the resulting interactions among τAgents. As seen in Figure 3, some agents may focus on data acquisition behaviors (e.g., continuous monitoring of health indicators), others on trajectory analysis (e.g., through machine learning algorithms), or patient trajectory anonymization and aggregation through a negotiation process (e.g., consent granting). We provide further details in [6] about how communication channels in τAgents use RDF as underlying representation model to achieve these interactions, embedded into standard agent protocols.

6 Real-time stream processing agents

The ubiquity of sensor technologies for monitoring and reporting health-related observations has already had a profound impact on clinical protocols, studies, and patient support services. The consumption and analysis of data streams produced by these sensors is challenging, especially for decentralized architectures. In [5] we propose a model for stream processing agents, capable of acting according to beliefs and goals represented as RDF streams, and including real-time constraints. The proposed architecture is centered on the notion of RDF Stream Processing *RSP agents*, which are autonomous entities as those in Section 5, with the added ability of communicating and exchanging RDF streams.

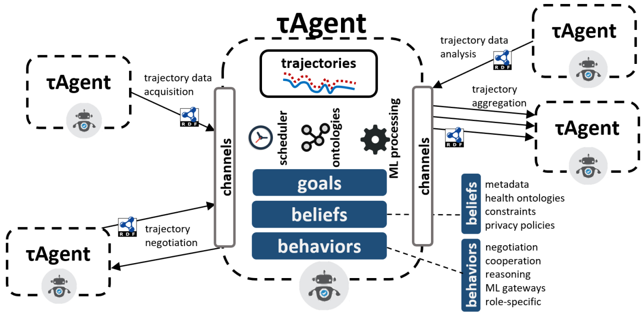

Fig. 3: Schematic view of τAgents for managing patient trajectories [6].

Following the Semantic Web principles, each agent and its resources are uniquely identified through URIs, and are equipped with endpoints that can be used to reach the RSP agent resources. The resources of each RSP agent includes the metadata of the RDF streams it manages, as well as other information relative to them (i.e., background RDF datasets, RDF stream buffers, ontology TBoxes, and RDF constraint rules). Depending on the nature of the RSP agent, it may implement different types of stream processing mechanisms such as continuous query processing, complex event processing, and stream reasoning.

As an example, the Stream Receiver in Figure 4 is a special case of a RSP agent that is able to consume streaming messages arriving at its inbox. The producer agent, or Stream Sender, can push streaming data at a given rate, which can be negotiated between the two parties. All streaming RDF stream messages are semantically annotated, e.g., using the data model in Section 3.

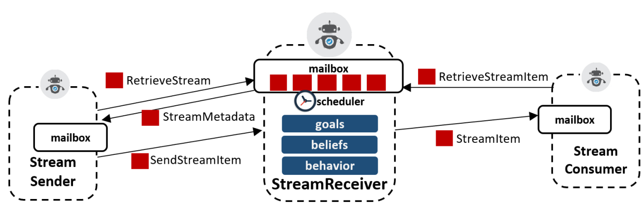

Fig. 4: RSP Stream Receiver interactions [5].

7 Privacy-aware Agents for Patient Data Exchange

Having described our approach for agent-based decentralized management of patient information, it would not be complete without considering complying with privacy protection regulations and guarantees. The enforcement of the General Data Protection Regulation (GDPR) [14] imposes, for instance, the need for explicit consent for data reuse, the right to timely receive all collected data, or the right to completely delete personal data. Beyond existing GDPR-compliant frameworks [2] that rely on centralized solutions, we explored and proposed in [7] the adoption of decentralized agent-based data privacy negotiation, coordination, and enforcement, using semantic representations of personal data privacy.

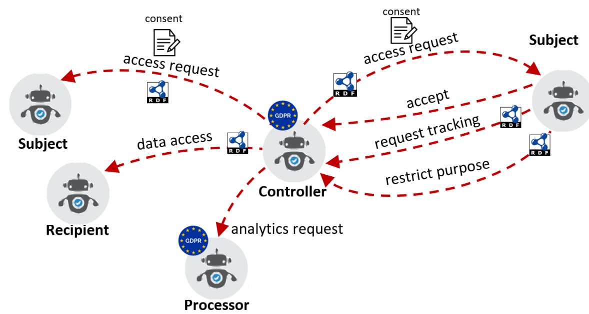

Fig. 5: Agent-based personal data privacy interactions [7].

We defined a set of minimal personal data privacy interaction requirements among agents and the design principles of privacy-aware agent interactions regarding personal data handling. Then, we integrated these concepts into an architecture design based on multi-agent protocol specifications, encoded as RDF messages. We base this specification in the Data Privacy Vocabulary (DPV) [13], developed by the Data Privacy Vocabularies and Controls Community Group of the W3C. We identified three main design principles to enable privacy compliant agent interactions:

- *Decentralized agents:* Following the nomenclature of the GDPR these privacy-aware agents can be: *data controllers*, *subjects*, *recipients*, and *processors* (Figure 5). Controllers refer to people or organizations that govern personal data processing. Subjects are the persons to which the data is related, while recipients are those to which personal information is disclosed. Processors are entities that process personal data on behalf of the controller.
- *Shared semantic vocabulary:* Semantic interoperability among these agents is dictated by the use of a common ontology for representing privacy data. We advocate the use of the Data Privacy Vocabulary (DPV)[2].
- *Data privacy agent interactions:* In principle, we base the definition of these interactions in existing FIPA protocols. For instance, a data consent request can be embedded in a request interaction protocol, or a data crowd-sourcing request can be represented as a ContractNet protocol.

8 Perspectives and Future Directions

Our vision for decentralized management of patient trajectory data using semantic agents has shown the potential of relying on autonomous processing and delegation of sensitive data processing. Nevertheless, it still remains necessary to take these building blocks and use them to implement our vision of decentralized

[2] https://www.w3.org/ns/dpv

patient trajectory management. In this section we provide a first glimpse of how this could be materialized.

As shown in Figure 6 we can represent the past events and circumstances of the patient history as a (retrospective) trajectory, in which each item is encoded as a semantically annotated entity (as described in Section 3). This trajectory is therefore part of a patient knowledge graph, in which —for example– a diagnosis can be represented as a `MedicalCondition`, with a specific ICD-10 code. The trajectory itself can also characterize the patient degree of distress/wellness, not only from a physical perspective, but also considering the social and psychological dimensions. For instance, the patient trajectory may exhibit a general physical decline after a surgery, while she can show general emotional improvement following physical therapy. If these trajectories are collected for a large number of participants, it is then possible to compute patterns and build models that allow predicting future outcomes. In Figure 6 we also show how predictive trajectories can be build using ML models, providing alternative trajectories depending on previous events, as well as on actions taken and achieved goals. As an example, given a certain retrospective trajectory, the system may predict three distinct trajectory outcomes depending on the degree of adherence to a therapy, enabling a personalized assessment of risks.

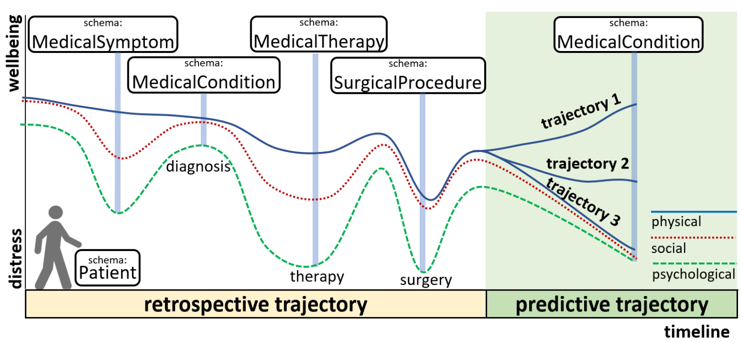

Fig. 6: Retrospective and predictive semantic trajectories.

Following our architectural and conceptual propositions described in Sections 4, 5, and 6, we provide a use-case example of how agent-based interactions can be used to enact trajectory-based analytics and support (Figure 7). Consider a cancer survivor who follows physical therapy to regain and maintain muscular strength. She is equipped with wearable sensors that collect motion data that can constitute part of her trajectory, using a stream-based model as explained in Section 6. However, this trajectory data is controlled by her through a patient trajectory agent (τAgent), and is not available to others unless explicit permissions are granted. As explained in Section 7, a coaching τAgent can emit a call for data, specifying purposes and other details through a consent, to which the patient agent can affirmatively respond. Once the access is granted, the data provided by the patient's sensor agents can be shared with the coaching agent. Similarly, retrospective trajectory data can be collected from a set of patient agents, and used to build a predictive trajectory model. The interaction between the coaching and processing τAgents allows the former to provide a personalized advice to the patient, while informing a therapist (through his corresponding agent) about the general progress, risks, and other relevant information.

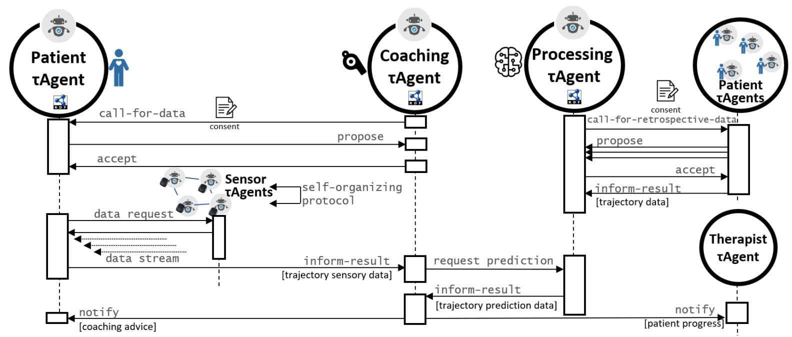

Fig. 7: Sequence of *tau*Agent interactions for the collection, processing and sharing of trajectory data.

Although our contributions in the works summarized in this paper provide preliminary evidence of applicability in the eHealth domain, there are still several research directions to be explored, included in our road-map. In the following we mention the most prominent ones:

(i) Trajectory modelling. Our proposed semantic model approach for patient trajectories is a first attempt that still needs to be validated in more complex scenarios, possibly in conjunction with data exchange, for instance in clinical trials where consent management and data restitution are critical. We also anticipate the need for domain-specific vocabularies/ontologies that may be required in order to better represent trajectory events, decision support, and delegation scenarios.

(ii) Implementation and validation. A key step in future research is the implementation and deployment of the proposed agent-based architecture, under real conditions and with a large cohort of patients. The complexity of the evaluation and validation of this approach is an additional challenge that will need to be addressed in order to provide an assessment from a technological and health-care point of view.

(iii) Decentralized explainability. A topic that has not been sufficiently explored concerns the explainability of trajectory analysis, and more so when ML algorithms are executed in a decentralized manner, with the risk of losing provenance information or missing relevant contextual information.

(iv) Trust and accountability. Unlike top-down approaches for data privacy compliance, our vision for decentralized personal data privacy interactions has the potential of allowing more flexible and scalable processing and protection of sensitive data. Nevertheless, it remains necessary to study the degrees of trust among participating agents, and to establish negotiation procedures when undesirable behaviors are detected. Accountable trajectory agents should be studied in order to address these concerns.

(v) Patient empowerment. Privacy protection is not the only aspect in which patients have the right to manage their own data. Patients should be provided with mechanisms for tracking their own data (e.g., for a clinical study) or even providing feedback, reusing it for her own benefit in a different context, or even exploring it through a set of third party services.

Acknowledgements

This work is partially supported by the H2020 project PERSIST: Patient-centered survivorship care plan after cancer treatment (GA 875406); and the HES-SO RCSO ISNet Project PERSA.

References

1. Alexander, G.L.: The nurse—patient trajectory framework. Studies in health technology and informatics **129**(Pt 2), 910 (2007)
2. Bourgeois, J., Kortuem, G., Kawsar, F.: Trusted and gdpr-compliant research with the internet of things. In: Proceedings of the 8th International Conference on the Internet of Things. pp. 1–8 (2018)
3. Calbimonte, J.P., Calvaresi, D., Dubosson, F., Schumacher, M.: Towards profile and domain modelling in agent-based applications for behavior change. In: Intl Conf on Practical Applications of Agents and Multi-Agent Sys. pp. 16–28 (2019)
4. Calbimonte, J.P., Dubosson, F., Schumacher, M.: Towards semantic models for proceeding and behavior change in ehealth applications. In: Proc. of the PD Track of the 15th Intl. Conference on Semantic Systems SEMANTiCS. pp. 41–45 (2019)
5. Calvaresi, D., Calbimonte, J.P.: Real-time compliant stream processing agents for physical rehabilitation. Sensors **20**(3), 746 (2020)
6. Calvaresi, D., Schumacher, M., Calbimonte, J.P.: Agent-based modeling for ontology-driven analysis of patient trajectories. J. Med. Systems **44**(9), 1–11 (2020)
7. Calvaresi, D., Schumacher, M., Calbimonte, J.P.: Personal data privacy semantics in multi-agent systems interactions. In: International Conference on Practical Applications of Agents and Multi-Agent Systems. pp. 55–67. Springer (2020)
8. Chung, J.Y., Lee, D.H., Park, J.H., Lee, M.K., Kang, D.W., Min, J., Kim, D.I., Jeong, D.H., Kim, N.K., Meyerhardt, J.A., et al.: Patterns of physical activity participation across the cancer trajectory in colorectal cancer survivors. Supportive Care in Cancer **21**(6), 1605–1612 (2013)
9. Guha, R.V., Brickley, D., Macbeth, S.: Schema. org: evolution of structured data on the web. Communications of the ACM **59**(2), 44–51 (2016)
10. Koutkias, V.G., Chouvarda, I., Triantafyllidis, A., Malousi, A., Giaglis, G.D., Maglaveras, N.: A personalized framework for medication treatment management in chronic care. IEEE Trans. on Inf. Tech. in Biomedicine **14**(2), 464–472 (2009)

11. Nelson, S.J., Schopen, M., Savage, A.G., Schulman, J.L.A., Arluk, N.: The mesh translation maintenance system: structure, interface design, and implementation. In: Medinfo. pp. 67–69 (2004)
12. Organization, W.H., et al.: Icd-10: international statistical classification of diseases and related health problems: tenth revision (2004)
13. Pandit, H.J., Polleres, A., Bos, B., Brennan, R., Bruegger, B., Ekaputra, F.J., Fernández, J.D., Hamed, R.G., Kiesling, E., Lizar, M., et al.: Creating a vocabulary for data privacy. In: OTM to Meaningful Internet Systems. pp. 714–730 (2019)
14. Voigt, P., Von dem Bussche, A.: The eu general data protection regulation (gdpr). A Practical Guide, 1st Ed., Cham: Springer International Publishing (2017)
15. Wu, H.S., Harden, J.K.: Symptom burden and quality of life in survivorship: a review of the literature. Cancer nursing **38**(1), E29–E54 (2015)