

Patient narratives – a still undervalued resource for healthcare improvement

Giovanni Spitale, Andrea Glässer, Mirriam Tyebally-Fang, Corine Mouton Dorey, Nikola Biller-Andorno

^a Institute of Biomedical Ethics and History of Medicine, University of Zurich, Switzerland

^b ZHAW School of Health Sciences, Institute of Public Health (IPH), Winterthur, Switzerland

Summary

In recent years, patient narratives have attracted increasing attention as a valuable source of insights into the subjective experience of healthcare. This paper outlines a best-practice approach to the collection, analysis, and use of patient narratives, based on current literature and on the experience of developing the Swiss Database of Individual Patient Experiences (DIPEX). The DIPEX project aims to provide a systematic and methodologically rigorous collection of patient narratives on various health situations and topics. This paper presents and details the DIPEX approach as a current standard in the field, offering a comprehensive overview and discussing the potential uses and benefits of patient narratives: improve healthcare practice, empower patients and caregivers, help structure better communication in healthcare, and contribute to medical teaching and learning.

Introduction

Research on patient narratives has steadily gained momentum over recent decades. For good reasons: patient narratives are unique windows on the subjective aspects of the healthcare experience – which is a relevant aspect of healthcare. Work on patient narratives can build empirical foundations for better patient-centred care, inform the development of patient-reported experience measures (PREMs) and of patient-reported outcome measures (PROMs), empower patients and caregivers, help structure better communication in healthcare, and contribute to medical teaching and learning. In a nutshell, narratives help us understand the very meaning of health, illness, and care for the key stakeholders of healthcare: patients [1]. Patient narratives can bring back people’s voices in patient-centred care; they are evidence, as in ‘evidence-based medicine’, complementary – not alternative – to quantitative evidence. Nevertheless, they need to be collected and studied with method and rigour to contribute to designing better healthcare.

Individual patient accounts that appear more or less randomly in the medical literature and media may be appealing and insightful but come with obvious limitations. They cannot capture variations, nor do they routinely provide detailed information about disease, diagnosis and treatment. A systematic collection, however, provides clear method-

ology and transparent analysis that is amenable to critical scrutiny. Results are contextualised and discussed, including other qualitative studies as well as quantitative work, to increase validity and reduce potential bias.

In this article, based on current literature and on our hands-on experience of developing a Swiss Database of Individual Patient Experiences (DIPEX), we detail a standardised, best-practice approach to the collection, analysis and use of patient narratives, in the form of semi-structured interviews – keeping in view that standards are currently evolving [2]. The Swiss DIPEX project aims to provide a systematic and methodologically rigorous collection of patient narratives on various health situations, such as diseases like dementia, chronic pain, Parkinson’s, COVID-19, multiple sclerosis, rare diseases and others, but also on health-related topics, such as intensive care management and coercion, pregnancy and prenatal testing, or possible future topics like adolescent obesity, risk of falls in older age, addictive behaviour among adolescents, or vocational reintegration at the workplace.

Background

The value of qualitative research has been long recognised in public health and health services research. The focus on methods that encourage participation and provide a deep insight into the subjective experience, which is one of the strengths of qualitative research, is now increasingly published and accepted in the mainstream medical literature: the Lancet commissions incorporate patient voices [3], and JAMA has a specific section on narratives [4]. The Equator Network, an international initiative seeking “to improve the reliability and value of published health research literature by promoting transparent and accurate reporting”, lists qualitative research as one of the main study types in health research [5–7].

Patient narratives can help close the gap between what really matters to the individual patient and what healthcare professionals perceive [8]. Beyond patient centredness, other dimensions of healthcare quality can be improved. Targeted care, based on patients’ needs and priorities, can be more effective and efficient, avoiding unnecessary expenses that do not add value for patients [9]. The saved resources can in turn be invested to deliver more timely and equitable care. Better understanding of patients’ perspectives contributes also to patient safety, for example,

Nikola Biller-Andorno
Institute of Biomedical
Ethics and History of Medi-
cine
University of Zurich
Winterthurerstrasse 30
CH-8006 Zurich
biller-andorno[at]
ibme.uzh.ch

through adapting information procedures to patients' emotional state and current cognitive receptivity. By sharing their experiences and by learning from others with similar conditions, patients and their relatives can feel supported and more able to cope with their disease [10]. Moreover, patient narratives can capture negative aspects of the interaction with the health care system, or with health care providers; therefore, the qualitative analysis of narrative material is a valuable strategy and a good source of insight for quality improvement in healthcare across different settings [11–13].

Patient narratives can also contribute to medical education and teaching, fostering understanding and improving communication. In fact, the creation of a database for a learning system has been identified as one of the core challenges that Swiss healthcare faces today [14].

Distant reading

To understand how patient narratives have been incorporated in healthcare scholarly work we adopted a 'distant reading' strategy [15, 16]. We used TopicTracker [17], a collection of programs written in Python to retrieve PubMed entries and to perform Natural Language Processing analyses on the corpus. Our query is fairly simple – but rather specific:

“1975/01/01” [Date –Publication]: “2021/12/31” [Date –Publication] AND (“patient s” [All Fields] OR “patients” [MeSH Terms] OR “patients” [All Fields] OR “patient” [All Fields] OR “patients s” [All Fields]) AND (“narration” [MeSH Terms] OR “narration” [All Fields] OR “narrative” [All Fields] OR “narratives” [All Fields] OR “narrative s” [All Fields] OR “narratively” [All Fields])

The query captures everything indexed in PubMed from 1975 (i.e., when the combination “patient narratives” appeared for the first time in a paper's keywords) to 2021.

Our analysis is focused (1) on normalised keywords and MeSH terms to describe the field and the main topics; and (2) on normalised journals (i.e.: the normalised count of

journals publishing this literature), to describe the impact, the target and the typical audience of these publications. Normalisation is performed in the same way for each entity, namely, normalised entity = count of entity / number of papers. The original dataset is available for replication and further exploration [18].

Our query captures a total of 26,739 papers (after duplicate removal) (fig. 1).

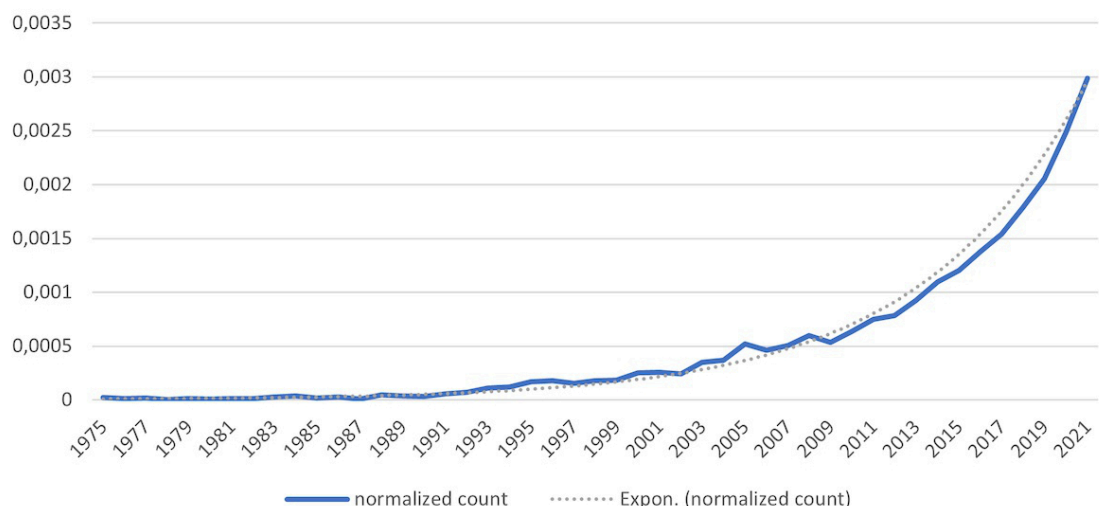
It is clear that the last decade saw the start of an exponential growth in publications dealing with “patient narratives”. Of note, in 2021 0.3% of all the literature indexed in PubMed touched this topic.

Keywords

The analysis of the most frequent 50 normalised keywords offers a first overview of what this field is about in terms of methodologies, conditions targeted and context of application of the findings. Keywords focus on:

1. The methodology and the approach, including analytical approach, empirical approach, philosophical approach, religious approach, qualitative research, natural language processing, patient experiences.
2. The target conditions, including death and euthanasia, mental health, cancer, genetics and reproduction, depression, dementia, stroke, pain / chronic pain, schizophrenia, diabetes, obesity, breast cancer. The target conditions remain rather stable over time.
3. The context in which this research can have a substantial impact, including professional patient relationship, palliative care, quality of life, nursing, rehabilitation, communication, narrative medicine, patient safety, primary care, diagnosis, surgery, prevention, medical education, and quality improvement. Some changes occurred over time in the “context” keywords, with some relevant new entries, becoming more popular in the last decade, namely: education, patient education, adherence. This suggests a gradual but steady switch of

Figure 1: Number of articles captured by our query on “patient narratives”, 1975–2021, normalised to total articles indexed in PubMed in the same year.



focus towards what we could call “patient-oriented research”.

MeSH terms

The analysis of the 50 most frequent normalised MeSH terms provides confirmation and further insight into what was highlighted by the keyword analysis. MeSH terms focus on:

1. The methodology and the approach, including narration, surveys and questionnaires, qualitative research, nursing methodology research, interviews, psychometrics, evaluation studies, retrospective studies, neuropsychological tests and randomised controlled trials.
2. The context in which this research can have a substantial impact, including interpersonal relations, attitude to health, self-concept, decision making, communication, medical history taking, attitude of health personnel, treatment outcome, social isolation, personal autonomy, verbal behaviour, social support, quality of life and grief. Of note, starting from 2011 “treatment outcome” became the most frequent “contextual” MeSH term in the corpus, followed by “quality of life” and “communication”, providing a clear indication of how patient narratives are mostly used now – informing the development of PROMs.

Journals

Journal trends suggest that although scholarly work on patient narratives tends to be published by discipline-specific journals (e.g. *Brain and Language*, *Social Science and Medicine*), it is gaining momentum and attention from top-notch medical journals, traditionally less inclined to publish this type of research (fig. 2). Although it is possible to speculate that the COVID-19 pandemic had a negative im-

act on the field, the most influential, high impact medical publications are accepting more research on patient narratives than before.

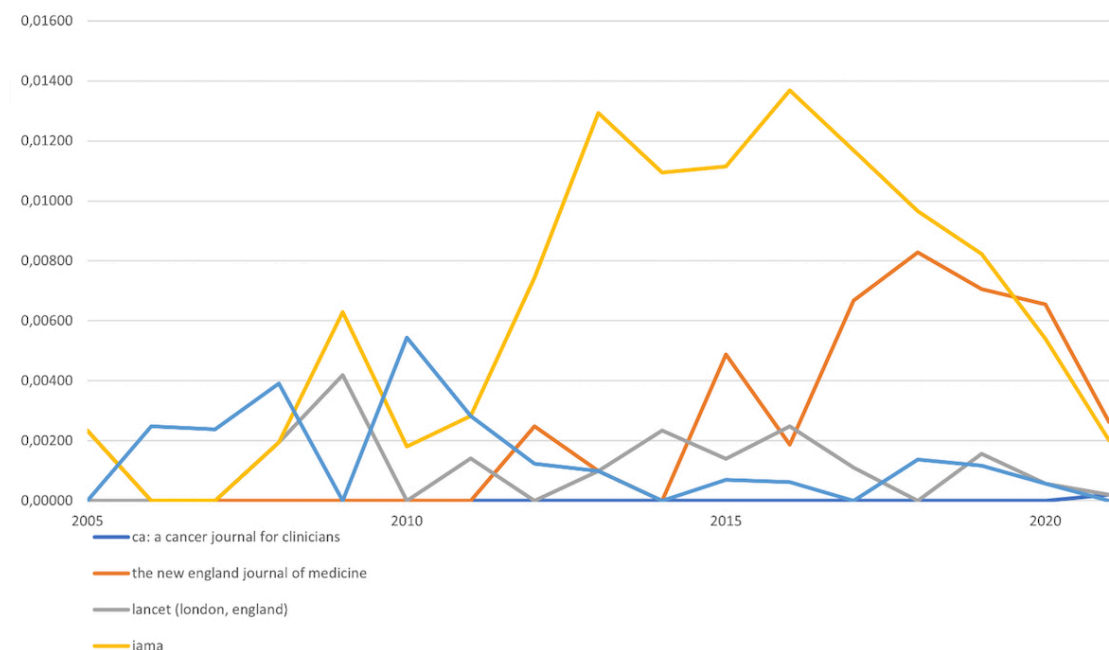
DIPEX International

Distant reading shows how participated and polyphonic the work on patient narratives has become, while maintaining some internal consistency. This richness is benefiting the field, but nevertheless we believe that some form of standardisation is needed. Standards ensure consistency, comparability, inter-compatibility and quality of the data, needs that have emerged even more strongly during the COVID-19 crisis: comparison of narrative data from different countries allowed study of the influence of structural factors like the healthcare system or response strategies on individual experiences, and hence, identification of good practices and improvement in quality of care.

Some work on defining solid standards for research on patient narratives, more specifically on patient interviews already exists, resulting from the joint efforts of the DIPEX international research community. DIPEX is a multi-media approach to collecting, analysing and disseminating patient interviews, which emerged in the early 2000s from the personal healthcare experience of Ann McPherson, a general practitioner, and Andrew Herxheimer, a clinical pharmacologist. The original idea of “a systematic collection and analysis of interviews with people about their experience of illness with evidence of the effects of treatments, and information about support groups and other resource materials” [19] grew over the years, as well as a research community.

DIPEX aims to “identify the questions that matter to people when they are ill” [19]. It collects interviews with patients with different diseases, which are made available on a website as audio, video or text, in accordance with the

Figure 2: Normalised trends for papers on “patient narratives” published in top medical journals, 2005–2021. Top medical journals have been identified and ordered by their reported impact factor in 2020: CA: A Cancer Journal for Clinicians – 120,83; the New England Journal of Medicine – 74,699; Lancet – 60,392; Journal of the American Medical Association – 51,273; Journal of Clinical Oncology – 32,956.



preferences of the interviewee. The idea is to inform patients, to educate healthcare professionals and to provide a “patient centred perspective to researchers and those who manage health services” [19].

The methodology has been further codified [20, 21], as well as the use and impact of the data for researchers, healthcare professionals, patients, and caregivers [10, 22–24].

In the last two decades, DIPEX international has grown into a solid network with research groups in 14 countries – including Switzerland [25].

The DIPEX.ch approach

In this section we present the current state of the art, the best practices and the main limitations of this kind of approach in developing our database of patient narratives, an open structure that allows the addition of further modules, providing data that can be used for future secondary analyses in other national or international research contexts, as done for COVID-19. In this way, the project and database aim to contribute not only to a smarter but to a wiser health care that is well anchored in patients’ needs, values and priorities.

Project definition

A DIPEX project starts with defining a “module”. Each module targets a specific health condition. The choice of the target depends on several factors, including evidence of knowledge gaps, suggestions by healthcare institutions, or requests from patient organisations.

Modules are conducted by a team of expert researchers, with experience in qualitative research, who have received specific formal training [21], and with a good command of multiple languages (a factor of crucial importance in a multilingual country such as Switzerland).

DIPEX Switzerland obtained a clarification of responsibility on the methodology from Zurich’s Cantonal Ethics Committee, which stated that “it does not fall within the scope of the Human Research Act” (BASEC-Nr. 2017-00678). This was later extended to the entire country, allowing us to “carry out the project throughout Switzerland without ethics committee approval” (BASEC-Nr. 2018-00050). New modules including substantial methodological changes are subject to a second round of ethical approval.

After a review of the available literature, the research team defines a research question and a topic-informed interview guide. This comprises an open narrative section and follow-up questions that allow understanding of specific topics, stemming from the research question. Each interview guide receives feedback from a module-specific advisory board, composed of healthcare professionals, other researchers, and patients. The feedback addresses both structural factors (e.g., appropriateness of the questions, salience) and linguistic factors (e.g., use of language-specific, non-translatable words, sentences or constructs). Our team of researchers comprises English, German, French and Italian speakers, with English being the common language. The interview guides are typically drafted in English, pilot-tested and subsequently translated. The translation is focused on maintaining an accurate interpretation of

the questions, as opposed to an accurate linguistic reproduction [26].

An example of interview guide is provided as appendix 1.

Data collection

As qualitative research typically aims for transferability and not representativity [27, 28], the sampling strategy follows a maximum variation approach [29]. Purposive sampling relies on two distinct notions: maximum variation and theoretical saturation. Maximum variation entails understanding which variables could influence the experience (age, gender, living arrangement, family background, condition-specific factors, ...) and representing these possible variations in the sample [19]. Importantly, these variables depend on the target condition. For example, “living arrangement” proved to be a very important factor in determining and shaping the experiences of COVID-19 patients, due to the transmissibility of the disease and to the quarantine/isolation mandates. Maximum variation is therefore defined before the data collection. Theoretical saturation is defined as the “point in data collection when no additional issues or insights are identified and data begin to repeat so that further data collection is redundant, signifying that an adequate sample size is reached.” [30]. Therefore, theoretical saturation is iteratively assessed while the data are being collected and analysed.

Participants are recruited through patients’ organisations, healthcare institutions, registries, although the strategy may vary depending on the target condition. The interviews are performed in the language participants are most comfortable in. When participants are not able to speak for themselves (e.g., due to dementia, or speech and cognitive disorders after a stroke), relatives are invited for an interview. In these cases, due consideration is given to specific ethical issues that may arise regarding informed consent procedures; authorisation and advice are sought from the local institutional review board (IRB). Interviews are conducted either in person or online, depending on the participant’s preference and context. Depending on the participant’s preferences, interviews are recorded as audio or video. Files are transcribed either manually or using GDPR (General Data Protection Regulation) compliant software [31], and proofread. Solid data protection protocols are of utmost importance in this phase, as the audio files may contain personal identifiers.

Coding and analysis

In order to determine theoretical saturation [30], the coding and analysis proceed in parallel with data collection. Anonymised transcripts are loaded in computer assisted qualitative data analysis software [32, 33], assigned the appropriate variables and coded. The preferential methodology is thematic analysis, a very flexible method “for identifying, analysing, and interpreting patterns of meaning (‘themes’) within qualitative data” [34, 35]. The approach is hybrid and iterative, with both top-down components (i.e., theory-driven codes, strongly connected with the research question and interview guide) and bottom-up components (i.e., data-driven codes, useful to capture emerging themes). Coding aims to identify, organise, describe and explain themes emerging in the corpus, coming to a shared

interpretation, and identifying opportunities for improving the quality and efficiency of healthcare.

Coding is a collaborative effort due to professional background and language. As the healthcare experiences encompasses medical, psychological and social factors, a multidisciplinary team can come to a more complete understanding of the corpus. To guarantee intersubjective consistency, every code is complemented by a memo, explaining its exact meaning and its intended use. When theoretical saturation is considered to have been reached, recruitment is stopped, and the coding work is consolidated and finalised.

To ensure that implicit meaning specific to a language are not lost, interviews are coded in the original language. The codes / coding tree remains in the common language of the research team. Particular words and phrases that are difficult to directly translate are highlighted by the researcher fluent in that language, and the cultural significance and meaning in the context of the interview transcript extensively memo-ed. To ensure all researchers have an understanding of the interview and an overview of the corpus of interview material, interview transcripts are translated into the common language using professional translation software. These translations are used only as a reference for the benefit of researchers outside of that language group and are not part of the formal coding and analysis process. Coded segments are translated to English only in the final step of preparing quotes for publication [36].

The goal is to enable the switch from “vertical reading” (i.e., traditional reading of the transcripts, line by line) to “horizontal reading” (i.e., retrieving all the fragments in which the participants talk about specific topics, defined by codes or groups of codes). This is the basis on which to go beyond individual experiences and understand topics, distilling the “collective experience” (i.e., the common traits), with an eye to the variations (i.e., the discrepancies). This mapping of themes emerging from the topics, their interplay and their interconnection informs the development of “OSOPs” (One Sheet of Paper), in which the evidence emerging from the narratives is contextualised, connected, and used to inform ethical reflection, either on specific topics or on the general management of the condition. This approach is based on thematic analysis, which in turn is derived from grounded theory [20].

Dissemination of the results

Once the analysis is finished, the data are prepared for dissemination and future use through a database. Based on the coding structure, the research team defines in the database a module-specific two-level taxonomy, cogent to report and portray the significant aspects of the “collective experience”. Each taxon is accompanied with a text describing the general lines of the content, again focussing both on trends and on variations. After completing the creation of the taxonomy, the research team identifies the most relevant coded fragments to attribute to the structure and creates in the database pseudonymised profiles for the participants from whose interviews the fragments come. The creation of “experiences” (i.e., selected, curated and classified fragments in which one interviewee tells a significant portion of their story) knits everything together. Experiences contain a title, a concise description, pseu-

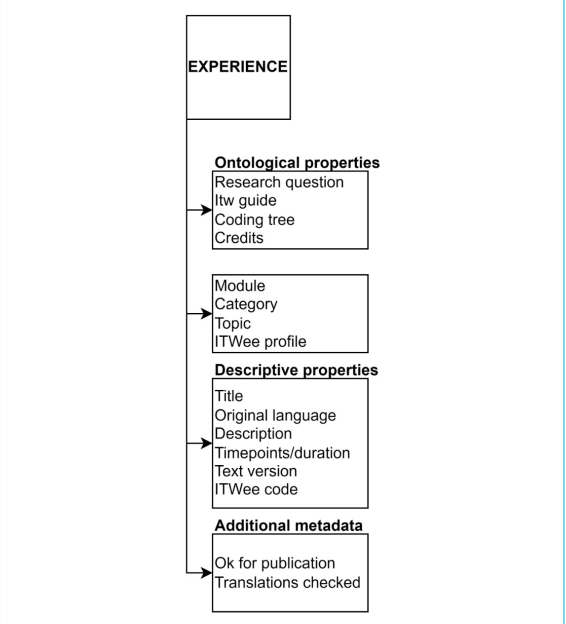
donymised transcript and (depending on the interviewee’s preferences) a link to the original audio or video. Experiences are thus the minimum viable entity of the database, and the database allows them to be stored (and retrieved) using their ontological properties, classification properties, descriptive properties or additional metadata. Properties are summarised in figure 3.

All the content of the database is automatically translated to English, German, French and Italian, and proofread. The database is a MariaDB system, hosted in the data centre of the University of Zurich, ensuring high data protection standards and adherence to FAIR (findable, accessible, interoperable, reusable) data principles.

Our first completed module contains a total of 334 experiences, belonging to 28 interviewees and taxonomised in 14 categories and 61 topics. The data are fed into the system using a dedicated interface and can be pulled using anything that supports MySQL queries. This allows multiple data-out interfaces, ranging from multimedia websites (as in the original DIPEX concept) to application programming interfaces (APIs). Although the website is mainly intended for lay users, APIs will allow a more specialistic use of the data (e.g., education, training, secondary research), for instance, defining ‘meta queries’ or allowing bulk download of tabular data. These functions maximise the (re)use potential of the dataset, increasing the value of the entire project and providing valuable resources for teaching, and current and future research lines.

Modules ready for the public are made available via www.dipex.ch and launched via a symposium including expert patients and healthcare professionals, providing excellent opportunities to present and discuss findings as done in Excellence in patient care in November 2021 [37]. The website is fully translated into German, French, Italian and English, and is constantly reviewed by means of qualitative and quantitative methods with a view to usability, accessibility and individual benefit, aiming to fulfil the requirements of the Health On the Net (HON) certificate [38], to which DIPEX subscribed. Participants can request

Figure 3: Properties of experiences in the DIPEX database.



to have their materials withdrawn from the website at any time.

Quality control and auditing

To ensure consistent quality of our data, alongside the application of the standards recognised by the Equator Network [6,7], we developed a quality control checklist specific to our processes, applied as a self-assessment tool for the researchers working on a module, and as the basis for peer review of the module. Quality control considers every step of the module production. The checklist is provided as supplementary material (appendix 2).

Data protection and data management

Interview data can contain sensitive personal information and are therefore to be treated with utmost care in terms of data protection. To protect our participants' privacy, our process comprises a two-level informed consent system and a set of data protection strategies.

Participants are required to sign the first informed consent before recording the interview. Signing this document, they declare they have read and understood the information sheet (appendix 3), that they were given the opportunity to ask questions and clarifications, that they agree to the interview being recorded (in video or audio), and that they consent to the use of the full pseudonymised interview for research and training. The first informed consent is provided as appendix 4.

After transcription and pseudonymisation, participants receive a copy of the transcript and a second informed consent form (appendix 5). They can allow the publication of the entire interview, of specific passages only or completely opt out from publication. Moreover, they can specify if they prefer to make their interview available as video, audio or text only.

If we need to send non-pseudonymised interviews to third parties (e.g., for transcription) we use SHA (Secure Hash Algorithm) encrypted containers and the third parties sign a nondisclosure agreement.

Upon data collection, identifiers are decoupled and an individual code is assigned to each participant. Data are stored in a password-protected folder in a server located at the University of Zurich, subject to incremental backup. The file containing the identities and contact details of the participants is SHA encrypted and stored on a different volume.

Data selected for publication are double-checked for complete pseudonymisation and saved in the database. The data-in interface has a two-factor multi-user authentication system, allowing granular permission management. Proof-reading can be done either directly via the data-in interface or exporting and re-importing translation files.

If a participant decides to opt-out from the study and requires data deletion, their code is retrieved, and the material attributed to that code is deleted from the server and from the database. Finally, the participant's entry is deleted from the key file.

Ethics

Any sustainable, future-oriented healthcare system needs to be conceived as a learning system. Learning, of course, requires feedback loops – such as patient experiences. Healthcare systems have been notoriously slow in developing smart ways of systematically and meaningfully taking patient experiences on board. For a long time, data collection was limited to “patient satisfaction”, focusing on items such as food or convenient parking. Patient experiences have evolved conceptually, covering outcome measures relevant to patients, but their collection has remained sketchy in many instances. Narrative data can complement this puzzle in a relevant way: by not imposing topics, approaches such as DIPEX can reveal what genuinely matters to patients and to what extent patient expectations are met by healthcare services. Listening to patients is indispensable if the goal is patient-oriented care. Translating insights into action is the next step that must follow. Researchers can contribute through the preparation of material (such as trigger films) based on patients' voices that can start inter-professional discussions of opportunities for improvement.

Another imperative regards the dissemination of findings. If participants dedicate their time to research, it is researchers' responsibility as good stewards of the entrusted data to ensure the results are accessible, not only to a narrow community of experts, but also to a broader range of academics and the interested public. Clearly, the use of data and results – particularly if re-identification cannot be excluded – is acceptable only with participants' written informed consent. Interviewees are not to be manipulated into revealing information they do not like to share. Treating participants with respect and heeding the limits they set is an obvious ethical rule that helps avoid exploitation.

Ethical issues may come up when participants reveal delicate information, such as suicidal ideations, drug abuse or involvement in illegal activities. In such cases, the complex interplay between the interviewer's duty to confidentiality, the participant's right to privacy and the safety of both is carefully assessed to define a course of action – e.g., putting the participant in contact with organisations or professionals that can offer qualified help.

Another caveat regards the groups and themes that receive scholarly attention. Some groups of patients may seem more attractive from a research point of view and more promising with a view to citations of future publications. However, from an ethical perspective, all groups of patients should have a chance to be heard, particularly those whose voice might be fainter, such as those who are marginalised, such as *sans-papiers*. At the same time, it is important to acknowledge that reaching these groups is not easy and often requires significant efforts. There is also an attention bias for certain health conditions, which may seem advantageous with a view to recruitment, funding or publication. This should not deter researchers from insisting on also covering conditions that are less easy to approach.

Finally, it is a matter of fairness to involve participants rather than just treating them as a data source. We recognise citizen science as a marker of good (open) science and therefore we try to include participants in every phase of

the process, from the definition and testing of the interview guide, to the analysis and interpretation of the results [39].

Limitations

Research on patient narratives is powerful and versatile, but it has some important limitations to keep in mind.

The approach we detailed is resource intensive. Researchers in charge of a module need solid training and must be very familiar with each aspect, from sample definition to interview techniques, to qualitative data analysis. Although some division of labour is possible, it is inadvisable to split a module in isolated and self-standing work packages. Some aspects can be automated: transcriptions and translations can be produced by software and proof-read by humans; we also automated the production and upload of the clips. This is an important innovation, as manual cutting and upload requires approximately 60–70 hours of human labour and about 7 hours of machine time per module, whereas programmatic cutting and uploading allows the same results to be achieved in about 8 hours of machine time. We are currently exploring the possibility to automate (to some extent) also the coding, but it is unclear whether it is possible to achieve a level of quality comparable to human coding [40]. In addition, after some years modules might require updates to remain current, adding new interviews and re-conducting the analysis.

This research can be rigorous and systematic. Nevertheless, personal bias can skew the coding and the interpretation. This risk can be mitigated by adopting a shared and multidisciplinary approach, by explaining codes with memos, and by exploring and challenging the preconceptions of the research team. Assessing theoretical saturation is an issue connected to this. On paper, theoretical saturation is considered reached when no new codes are emerging [41]. In practice, if the coding team did not approach coding with enough rigour, it could declare theoretical saturation without reaching it. Because of the role of theoretical saturation in assessing the rigor of qualitative research, both its definition and the assessment process should be defined in a transparent way for each sample [30].

Results from research on patient narratives are particularly vulnerable to “data torturing”. They make sense when considered in their unity, complexity and context; to build reflections on the entire movie, rather than on a snapshot, hosting a DIPEX project in a research unit natively offering interdisciplinary expertise (such as a biomedical ethics unit) is a good mitigation strategy.

Finally, results must be operationalised and integrated into patient care. This requires a certain predisposition to integrate this evidence on the part of the healthcare system – which currently can depend a lot on personal sensitivities and inclinations.

Conclusion

When compared with other current and emerging collections of patient narratives (e.g., social media pages/groups), the approach we detailed offers significant advantages: conflict of interest is assessed and avoided; data collection and analysis are rigorous and systematic; public-facing content is carefully curated, and the entire process is supported by a lively research community.

A recent systematic review on risks and benefits of patient narratives concluded that “patient narratives seem to be a promising means to support users in improving their understanding of certain health conditions and possibly to provide emotional support and have an impact on behavioural changes” [23]. This approach goes beyond informing the development of evidence-based PROMs and PREMs, decision aids or trigger films. Our data structure is built with an eye to the future: curated collections hosted on FAIR databases will provide a fundamental infrastructure for natural language processing approaches to patient narratives, for the training of conversational artificial intelligence systems, for medical education, and for data sharing in the context of international studies.

Finally, this approach can help integrate the contributions of different disciplines sitting at the interdisciplinary table of biomedical ethics – medicine, nursing studies, philosophy, sociology, psychology, anthropology, and others – in the interest of a joint goal: patient-centred care. At the same time, making patients’ stories publicly available endorses important cultural changes currently under way, reducing knowledge and power asymmetries between healthcare practitioners and patients (and their relatives), and fostering genuinely personalised healthcare by providing opportunities to listen to what matters to patients.

Acknowledgements

For their contribution in developing DIPEX modules, collecting and analysing data (in alphabetical order): Yolanda Chacón Gámez; Joana Cuenoud; Daniel Drewniak; Andrea Durisch; Beatrix Göcking; Many Hendriks; Martina Hodel; Yvonne Ilg; Susanne Jöbges; Silvia Lazarotto; Andrea Lehner; Anke Maatz; Tania Manríquez Roa; Joelle Ott; Andrea Radvanszky; Johann Roduit; Mengzhen Schwarz; Bettina Schwind; Karin Seiler; Nina Streeck; Sebastian Wäscher; Henrike Wiemer; Kristina Würth.

The organisations that support DIPEX.ch and its research (in alphabetical order): Alzheimer Schweiz; Collegium Helveticum; Digitalisierungsinitiative des Kantons Zürich DIZH; Swiss Academy of Medical Sciences; Schweizerische Multiple Sklerose Gesellschaft; Swiss Cancer Research Foundation (KFS-4690-02-2019); University of Zurich; Zurich University of Applied Sciences, School of Health Professionals.

For their support in starting this ambitious research platform, the DIPEX International group and, in particular, Gabriele Lucius-Hoene, Martina Breuning and Christine Holmberg.

The DIPEX Advisory Board for its valuable input and Dr Hermann Amstad for strategic advice.

Potential competing interests

All authors have completed and submitted the International Committee of Medical Journal Editors form for disclosure of potential conflicts of interest. No potential conflict of interest was disclosed.

References

- Pallai E, Tran K. Narrative Health: Using Story to Explore Definitions of Health and Address Bias in Health Care. *Perm J*. 2019;23(1):18–052. <http://dx.doi.org/https://doi.org/> <http://dx.doi.org/10.7812/TPP/18-052>.
- Deutsches Netzwerk Gesundheitskompetenz. *Erfahrungsberichte im Gesundheitswesen*. Dtsch. Netz. Gesundheitskompetenz DNGK. 2022. <https://dngk.de/erfahrungsberichte/>.
- Han MK, Martinez CH, Au DH, Bourbeau J, Boyd CM, Branson R, et al. Meeting the challenge of COPD care delivery in the USA: a multi-provider perspective. *Lancet Respir Med*. 2016;4(6):473–526. <http://dx.doi.org/https://doi.org/> [http://dx.doi.org/10.1016/S2213-2600\(16\)00094-1](http://dx.doi.org/10.1016/S2213-2600(16)00094-1).
- JAMA. Narrative Medicine. *jamanetwork.com*. 2022. <https://jamanetwork.com/collections/5766/narrative-medicine>.
- Equator Network. About us | The EQUATOR Network. *Equat.-Network.org*. 2022. <https://www.equator-network.org/about-us/>.

6. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med.* 2014 Sep;89(9):1245–51. <http://dx.doi.org/https://doi.org/10.1097/ACM.0000000000000388>.
7. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care.* 2007 Dec;19(6):349–57. <http://dx.doi.org/https://doi.org/10.1093/intqhc/mzm042>.
8. Kleinman A. *The Illness Narratives: Suffering, Healing, And The Human Condition.* New York: Basic Books; 1989.
9. Porter ME. What is value in health care? *N Engl J Med.* 2010 Dec;363(26):2477–81. <http://dx.doi.org/https://doi.org/10.1056/NEJMp1011024>.
10. Ziebland S, Lavie-Ajayi M, Lucius-Hoene G. The role of the Internet for people with chronic pain: examples from the DIPEX International Project. *Br J Pain.* 2015 Feb;9(1):62–4. <http://dx.doi.org/https://doi.org/10.1177/2049463714555438>.
11. Tsianakas V, Maben J, Wiseman T, Robert G, Richardson A, Madden P, et al. Using patients' experiences to identify priorities for quality improvement in breast cancer care: patient narratives, surveys or both? *BMC Health Serv Res.* 2012 Aug;12(1):271. <http://dx.doi.org/https://doi.org/10.1186/1472-6963-12-271>.
12. Brand G, Osborne A, Wise S, Isaac C, Etherton-Bear C. Using MRI art, poetry, photography and patient narratives to bridge clinical and human experiences of stroke recovery. *Med Humanit.* 2020 Sep;46(3):243–9. <http://dx.doi.org/https://doi.org/10.1136/medhum-2018-011623>.
13. Locock L, Kirkpatrick S, Brading L, Sturmeijer G, Cornwell J, Churchill N, et al. Involving service users in the qualitative analysis of patient narratives to support healthcare quality improvement. *Res Involv Engagem.* 2019 Jan;5(1):1. <http://dx.doi.org/https://doi.org/10.1186/s40900-018-0133-z>.
14. Biller-Andorno N, Zeltner T. Individual Responsibility and Community Solidarity—The Swiss Health Care System. *N Engl J Med.* 2015 Dec;373(23):2193–7. <http://dx.doi.org/https://doi.org/10.1056/NEJMp1508256>.
15. Franzini G, Jänicke S, Scheuermann G, et al. On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. *A State-of-the-Art (STAR) Report.* 2015.
16. Moretti F. *Distant Reading.* 2013.
17. Spitale G, Biller-Andorno N. TopicTracker: a Python pipeline to search, download and explore PubMed entries. 2021. <https://doi.org/http://dx.doi.org/10.5281/zenodo.4643876>.
18. Spitale G. PubMed literature on patient narratives 1975–2021. 2022. <https://doi.org/http://dx.doi.org/10.5281/zenodo.6534889>.
19. Herxheimer A, McPherson A, Miller R, Shepperd S, Yaphe J, Ziebland S. Database of patients' experiences (DIPEX): a multi-media approach to sharing experiences and information. *Lancet.* 2000;355(9214):1540–3. [https://doi.org/http://dx.doi.org/http://dx.doi.org/10.1016/S0140-6736\(00\)02174-7](https://doi.org/http://dx.doi.org/http://dx.doi.org/10.1016/S0140-6736(00)02174-7).
20. Ziebland S, McPherson A. Making sense of qualitative data analysis: an introduction with illustrations from DIPEX (personal experiences of health and illness). *Med Educ.* 2006 ;40(5):405–14. <https://doi.org/http://dx.doi.org/http://dx.doi.org/10.1111/j.1365-2929.2006.02467.x>.
21. Medical Sociology and Health Experiences Research Group (MS & HERG). *Researcher's Handbook – Healthtalk Modules.* 2020.
22. Ziebland S, Herxheimer A. How patients' experiences contribute to decision making: illustrations from DIPEX (personal experiences of health and illness). *J Nurs Manag.* 2008;16(4):433–9. <https://doi.org/http://dx.doi.org/http://dx.doi.org/10.1111/j.1365-2834.2008.00863.x>.
23. Drewniak D, Glässel A, Hodel M, Biller-Andorno N. Risks and Benefits of Web-Based Patient Narratives: systematic Review. *J Med Internet Res.* 2020 ;22(3):e15772. <https://doi.org/http://dx.doi.org/http://dx.doi.org/10.2196/15772>.
24. Christensen V, Parker K, Cottrell E. Leveraging a qualitative data repository to integrate patient and caregiver perspectives into clinical research. *J Clin Transl Sci.* 2021 ;5(1):e155. <https://doi.org/http://dx.doi.org/http://dx.doi.org/10.1017/cts.2021.822>.
25. Ziebland S, Grob R, Schlesinger M. Polyphonic perspectives on health and care: reflections from two decades of the DIPEX project. *J Health Serv Res Policy.* 2021 ;26(2):133–40. <https://doi.org/http://dx.doi.org/http://dx.doi.org/10.1177/1355819620948909>.
26. Schaffner C. From "Good" to "Functionally Appropriate": Assessing Translation Quality. *Curr Issues Lang Soc.* 1997;4(1):1–5. <http://dx.doi.org/10.1080/13520529709615476>.
27. Morse JM. Determining Sample Size. *Qual Health Res.* 2000;10(1):3–5. <https://doi.org/http://dx.doi.org/http://dx.doi.org/10.1177/104973200129118183>.
28. Marshall MN. Sampling for qualitative research. *Fam Pract.* 1996 ;13(6):522–5. <https://doi.org/http://dx.doi.org/http://dx.doi.org/10.1093/fampra/13.6.522>.
29. Coyne IT. Sampling in qualitative research. Purposeful and theoretical sampling; merging or clear boundaries? *J Adv Nurs.* 1997;26(3):623–30. <https://doi.org/http://dx.doi.org/http://dx.doi.org/10.1046/j.1365-2648.1997.t01-25-00999.x>.
30. Hennink MM, Kaiser BN. *SAGE Research Methods Foundations.* 2020. <https://doi.org/http://dx.doi.org/http://dx.doi.org/10.4135/9781526421036822322>.
31. HappyScribe. Security and Data Protection. www.happyscribe.com. 2022. <https://www.happyscribe.com/security>.
32. Kuckartz U, Rädiker S. Analyzing qualitative data with MAXQDA: text, audio, and video. 2019. <https://doi.org/http://dx.doi.org/http://dx.doi.org/10.1007/978-3-030-15671-8>.
33. Woolf NH, Silver C. Qualitative Analysis Using MAXQDA: The Five-Level QDATM Method; 2017. <https://doi.org/http://dx.doi.org/http://dx.doi.org/10.4324/9781315268569>.
34. Braun V, Clarke V, Hayfield N, et al. In: Liampittong P, editor. *Handbook of Research Methods in Health Social Sciences.* 2018. pp. 1–18. https://doi.org/http://dx.doi.org/http://dx.doi.org/10.1007/978-981-10-2779-6_103-1.
35. Braun V, Clarke V. In: *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological.* Washington, DC, US: American Psychological Association 2012. 57–71. <https://doi.org/http://dx.doi.org/http://dx.doi.org/10.1037/13620-004>.
36. Ferguson G, Pérez-Llantada C, Plo R. English as an international language of scientific publication: a study of attitudes. *World Engl.* 2011;30(1):41–59. <https://doi.org/http://dx.doi.org/http://dx.doi.org/10.1111/j.1467-971X.2010.01656.x>.
37. IBME. Excellence in Patient Care Symposium 2021 – The Swiss Database of Individual Patient Experiences (DIPEX.ch). 2021. <http://www.ibme.uzh.ch/en/Biomedical-Ethics/Agenda/Previous-Events/Excellence-in-Patient-Care-Symposium-2021.html>.
38. HON. Our commitment to reliable health and medical information. 2017. <https://www.hon.ch/HONcode/Patients/Visitor/visitor.html>.
39. Participatory Science Academy. What is participatory research? 2020. <http://www.pwa.uzh.ch/en/aboutus/What-is-participatory-research-.html>.
40. Spitale G, Biller-Andorno N, Germani F. Concerns Around Opposition to the Green Pass in Italy: Social Listening Analysis by Using a Mixed Methods Approach. *J Med Internet Res.* 2022 ;24(2):e34385. <https://doi.org/http://dx.doi.org/http://dx.doi.org/http://dx.doi.org/10.2196/34385>.
41. Nascimento L CN, Souza TV de, Oliveira IC dos S, et al. Theoretical saturation in qualitative research: an experience report in interview with schoolchildren. *Rev Bras Enferm* 2018;71:228–33. <https://doi.org/http://dx.doi.org/http://dx.doi.org/http://dx.doi.org/10.1590/0034-7167-2016-0616>.

Appendices

The appendices can be downloaded as separate files at <https://doi.org/10.57187/smw.2022.40022>.