Review

Gynecology Meets Big Data in the Disruptive Innovation Medical Era: State-of-Art and Future Prospects

Rola Khamisy-Farah [1], Leonardo B. Furstenau [2], Jude Dzevela Kong [3], Jianhong Wu [3] and Nicola Luigi Bragazzi [3,*]

1 Clalit Health Service, Akko, Azrieli Faculty of Medicine, Bar-Ilan University, Safed 13100, Israel;
   rkhamisy@yahoo.com
2 Department of Industrial Engineering, Federal University of Rio Grande do Sul,
   Porto Alegre 90035-190, Brazil; leonardo.furstenau@ufrgs.br
3 Laboratory for Industrial and Applied Mathematics (LIAM), Department of Mathematics and Statistics,
   York University, Toronto, ON M3J 1P3, Canada; jdkong@yorku.ca (J.D.K.); wujhhida@gmail.com (J.W.)
* Correspondence: bragazzi@yorku.ca

Abstract: Tremendous scientific and technological achievements have been revolutionizing the current medical era, changing the way in which physicians practice their profession and deliver healthcare provisions. This is due to the convergence of various advancements related to digitalization and the use of information and communication technologies (ICTs)—ranging from the internet of things (IoT) and the internet of medical things (IoMT) to the fields of robotics, virtual and augmented reality, and massively parallel and cloud computing. Further progress has been made in the fields of addictive manufacturing and three-dimensional (3D) printing, sophisticated statistical tools such as big data visualization and analytics (BDVA) and artificial intelligence (AI), the use of mobile and smartphone applications (apps), remote monitoring and wearable sensors, and e-learning, among others. Within this new conceptual framework, big data represents a massive set of data characterized by different properties and features. These can be categorized both from a quantitative and qualitative standpoint, and include data generated from wet-lab and microarrays (molecular big data), databases and registries (clinical/computational big data), imaging techniques (such as radiomics, imaging big data) and web searches (the so-called infodemiology, digital big data). The present review aims to show how big and smart data can revolutionize gynecology by shedding light on female reproductive health, both in terms of physiology and pathophysiology. More specifically, they appear to have potential uses in the field of gynecology to increase its accuracy and precision, stratify patients, provide opportunities for personalized treatment options rather than delivering a package of "one-size-fits-it-all" healthcare management provisions, and enhance its effectiveness at each stage (health promotion, prevention, diagnosis, prognosis, and therapeutics).

Keywords: big data; fast and smart data; disruptive innovation medical era; gynecology

1. Theoretical Background: The Current Medical Era as the "Disruptive Innovation Era"

Tremendous scientific and technological achievements have been revolutionizing the current medical era, shaping the way in which physicians practice their profession and deliver healthcare provisions. This is due to the convergence of various advancements related to digitalization and the use of information and communication technologies (ICTs)—ranging from the internet of things (IoT) and the internet of medical things (IoMT) to the fields of robotics, virtual and augmented reality, and massively parallel and cloud computing. Further progress has been made in the fields of addictive manufacturing and three-dimensional (3D) printing, sophisticated statistical tools such as big data visualization and analytics (BDVA) and artificial intelligence (AI), the use of mobile and smartphone applications (apps), remote monitoring and wearable sensors, and e-learning, among others.

Other drivers of and catalysts for innovation are represented by developments in the fields of P6 medicine (where the 6 Ps stand for personalized, predictive, preventive, participatory, psycho-cognitive, and public), nanomedicine, molecular medicine, and nanotherapy—including the use of new nano(bio)technological materials, and genetic, genomic, and post-genomic engineering and editing [1].

All these profound changes characterize the so-called "disruptive era" (Figure 1) [2], which is dramatically impacting all medical specializations—including gynecology. As such, gynecology has recently branched out into various sub-specialties, like:

- the internet of gynecological things.
- eGynecology (from the combination of gynecology and eHealth, which, in turn, is an abbreviation for electronic health, conceived as the use of electronic devices and tools aimed at improving and enhancing health-related outcomes) [3].
- mGynecology (from the combination of gynecology and mHealth, which, in turn, is an abbreviation for mobile health, that is to say the use of mobile and wireless technologies to achieve health-related purposes and objectives).
- uGynecology (from the combination of gynecology and uHealth, which, in turn, is an abbreviation for ubiquitous health, or the application of wearable sensors to constantly monitor in real-time the health status of individuals).
- tele-gynecology (from the combination of gynecology and tele-health/tele-medicine, which refers to the innovative and emerging practice of delivering healthcare provisions remotely).
- nano-gynecology (from the combination of gynecology and nano-medicine, which is the practice of the medicine at the nano-scale level) [4].
- precision and predictive gynecology (which implies the customizing, tailoring, and individualizing of healthcare provisions related to gynecology, based on the individual features of each patient).
- big and smart gynecology.
- 3D gynecology (which is the application of bioengineering and additive manufacturing processes to deposit bio-active materials, known as bio-inks, in order to develop and create tissue-like structures for regenerative or other medical purposes) [5].

Within this new conceptual framework [6], big data represents a massive set of data, characterized by different properties and features. These can be categorized both from a quantitative and qualitative standpoint, and include:

1. velocity (so-called "fast data", which indicates the speed at which big data can be generated, released, and analyzed, almost in real-time, enabling data prediction, such as nowcasting or forecasting).
2. variety (because they can be produced by various channels and sources).
3. variability (since they are changeable and constantly in flux).
4. volume (related to the amount and magnitude of available data and information).
5. veracity/verification (related to data trustworthiness, accuracy, and reliability).
6. visualization (related to readability, affordability, and accessibility of data).
7. visibility (referring to the state of being able to be seen, this state is also known as "big visibility").
8. virality (the speed at which data and related information spread along a people to people (P2P) network).
9. viscosity (or data friction or resistance, which refers to the heterogeneity of particularly complex data, which may be difficult to incorporate, integrate and combine).
10. virtual (referring to the digital/computational nature of big data).
11. value (in terms of advantages and benefits deriving from the collection, use, and analysis of big data). This implies that, after processing and analysis, big data can become smart and insightful Data.

Compared to other kinds of data, big data are difficult to store, pre-process, and analyze in an efficient way by utilizing classical techniques, thus requiring new computational resources, infrastructures, and ad hoc facilities.
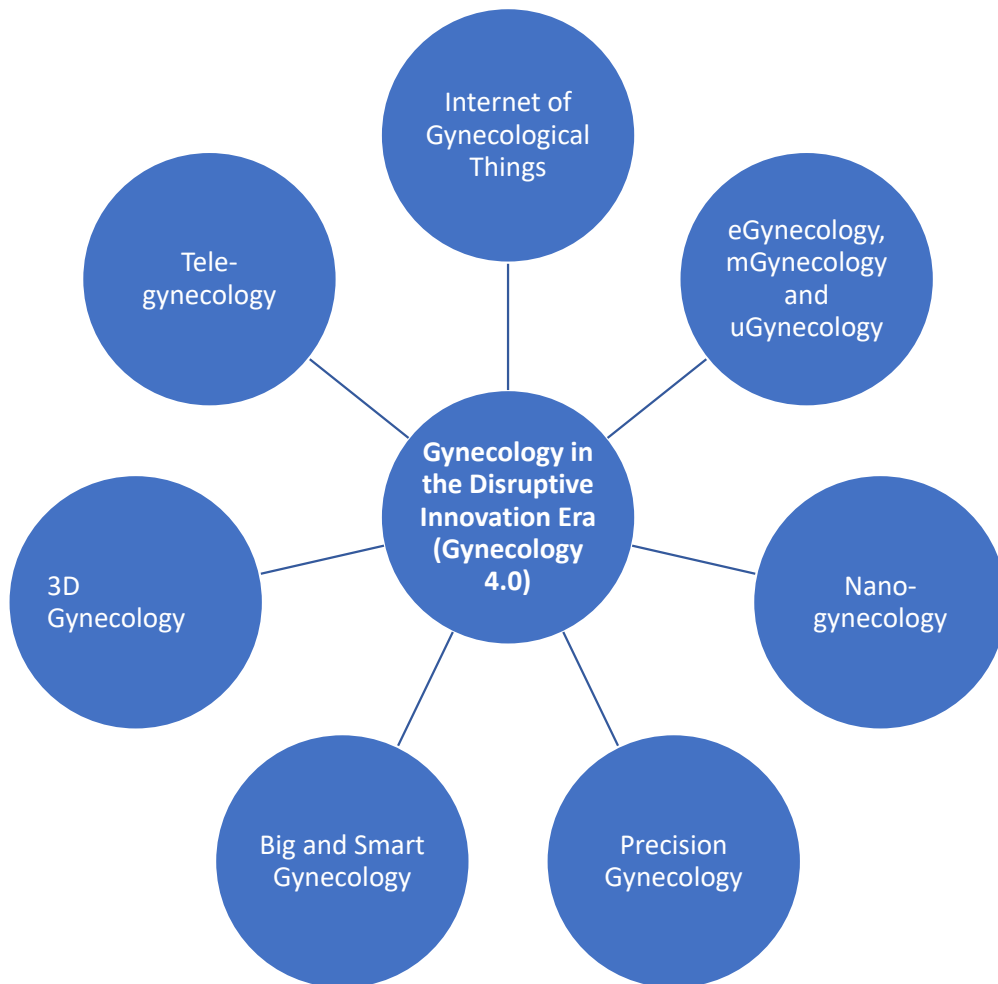

Figure 1. The convergence of advancements and progress from different fields (information and communication technologies, digitalization, molecular biology, biochemistry, biophysics, nano-engineering, and nano(bio)technology).

Genomics, post-genomics (for instance, proteomics and metabonomics/metabolomics, among others) and biocomputational (bioinformatics or cheminformatics) repositories, clinical registries, imaging techniques and wearable sensors, and databases consulted by patients as well as by the lay people represent data streams producing molecular, clinical, imaging, and digital big data, respectively (as summarized in Table 1) [6,7].

Table 1. An overview of the different kinds of Big Data.

Type of Big Data	Sources
Molecular Big Data	Wet-lab, microarrays
	Bioinformatics/cheminformatics repositories
Computational/Clinical Big Data	Electronic Health Records (EHRs) and clinical databases
Imaging Big Data	Wearable sensors, imaging approaches
Digital Big Data	Website searches

2. The Aims of the Present Integrative Review

This comprehensive, integrative review aims at providing readers with an updated overview of the challenges and future research directions of big and smart data applied

in the field of gynecology. The different types of big data will be presented, together with some of their major examples and applications for enhancing human female fertility, reproductivity, and treating gynecological diseases. Ethical implications will also be envisaged.

3. Gynecology and OMICS Data

Diseases result from the complex, multi-factorial interplay between a variety of exogenous and endogenous factors, including biological/individual, behavioral (lifestyles, such as smoking, dietary habits, and physical activity), and environmental variables. As such, they have unique molecular signatures, in terms of prognosis, response to pharmacological agents, co-morbidity, mortality, and survival rates. In this sense, the same disease may be considered as an umbrella term, collecting various pathophysiological and clinical manifestations. Deep phenotyping (quantitative histo-pathology/cyto-pathology) and clustering techniques enable us to differentiate and stratify patients in terms of outcomes and prognosis, such as response to drugs, surgery, or other treatment options, thus potentially guiding and informing the decision-making process according to available evidence [8,9].

Genomics studies can help identify a panel of genetic biomarkers, aimed at early detection of gynecological disorders (such as malignancies) and curb their morbidity/mortality rates. For instance, for ovarian cancer, a panel made up of cancer antigen 125 (CA-125), human epididymis protein 4 (HE4), cancer antigen 19-9 (CA 19-9), epidermal growth factor receptor (EGFR), granulocyte colony-stimulating factor (G-CSF), eotaxin, interleukin-2 receptor (IL-2R), circulating vascular cell adhesion molecule-1 (cVCAM), and macrophage migration inhibiting factor (MIF) seems to exhibit a satisfactory specificity and sensitivity [10]. The Gene Expression Predictor of High-grade Serous Ovarian Carcinoma Molecular SubTYPE (PrOTYPE) is another example of genomic panel comprising of 55 genes that has been developed and validated for predictive purposes in ovarian cancer patients, utilizing published gene expression data from 1650 cancers. Analyses have shown a statistically significant association between a set of variables, including age, residual disease, tumor stage, and tumor-infiltrating lymphocytes, as well as other relevant clinical-pathological outcomes and biological parameters, with the panel exhibiting an accuracy greater than 95% [11].

Similar panels can be devised at the proteomics level (see, for instance [12–15]). For example, Timms et al. [15] exploited different complementary profiling platforms and strategies to identify biomarker candidates in a statistically robust fashion. Authors were able to pin down four genes (namely, alpha-1 antitrypsin or A1AT, secretory leukocyte peptidase inhibitor or SLPI, apolipoprotein A-IV or APOA4, and vitamin D binding protein or VDBP) that could significantly differentiate between malignant and benign cases. On the other hand, either alone or even when combined, these biomarkers failed to overperform with respect to the "classical" biomarker CA-125, which remains the gold standard in terms of accuracy for diagnosing ovarian cancer. However, SLPI seemed to act as a satisfactory early biomarker, even though data were too preliminary and no firm conclusion could be drawn, thus warranting further research.

Furthermore, new, emerging biomarkers have been discovered in the fields of genomics and proteomics, like exosomes, which are involved in the cellular cross-talks and related biological processes [16]. Nutrigenomics [17,18], epigenomics [19] and nutriepigenomics [20,21] investigate how dietary habits and nutrition, as well as environmental factors, and their complex, non-linear interactions can finely modulate the human genome and its expression, and impact on human diseases, including gynecological conditions. For instance, a study explored the relationship between folic acid supplementation and one carbon metabolism gene single nucleotide polymorphisms (SNPs) on small-for-gestational-age (SGA) births [17]. Other studies assessed the association between pregnancy and childhood epigenetics [19,20].

Besides proteomics, other post-genomics specialties include transcriptomics [22] and cytomics [22–24], which enable the study of gene transcripts and biochemistry and bio-

Int. J. Environ. Res. Public Health **2021**, *18*, 5058

5 of 13

physics at the single cellular level, respectively. Another emerging OMICS discipline aims at exploring and dissecting the molecular basis of the vaginal microbiome, conceived as a specific compartment of the human microbiome [25]. The vagina is characterized by specific features, such as the presence of few bacteria, in contrast to other parts of the human body. These microbes include glycogen-using, lactic acid-producing, estrogen-dependent *lactobacilli*, and other fermentative microbial species, which, together with epithelial cells, are responsible for the acid vaginal pH [26].

Various vaginotypes exist based on the composition of vaginal microbiota, which can have an impact on vaginal dysbiosis, vaginosis, sexually transmitted diseases (STDs), such as the infections by human papillomavirus (HPV), and birth and pregnancy outcomes [27–33]. As such, 'vaginal and maternal microbiomics' is an emerging field of super-specialties within microbiomics, which appears promising in terms of predictive power. This paves the way for tailored therapeutics—the so-called pharmacomicrobiomics and pharmacoculturomics [34], that is to say, drugs designed and developed based on our knowledge of molecular interactions between the human vaginal microbiome and potential pharmacological agents.

Metabolomics/metabonomics [35], exosome metabolomics (exometabolomics) [35], and multi-OMICS approaches [36,37] represent further innovative strategies for (sub-)typing gynecological patients. Finally, phenomics [38] aims at systematically and comprehensively measuring and quantitatively assessing the human phenome, which is the array of physical, chemical, biochemical, and biophysical traits expressed and/or generated by an individual during its development, adaptation to genetic influences, and environmental stimuli. For instance, in Denmark, Westergaard et al. [38] carried out a phenome-wide study to explore the molecular basis of recurrent pregnancy loss (RPL) utilizing a cohort of 1,370,896 never-pregnant women aged 12–40 years in the study period between 1977 and 2016. The authors used a population-wide healthcare registry and found that RPL represented a major risk factor for a wide array of disorders, including cardiovascular, autoimmune, and mental health diseases. This could shed light on the etiopathogenetic mechanisms underlying RPL and provide new management and treatment options, warranting further research in the field. Other phenomics studies have explored gynecological disorders, such as polycystic ovary syndrome [39] and reproductive issues [40–44].

All these OMICS disciplines and specialties are briefly overviewed in Table 2.

Table 2. An overview of the main OMICS disciplines useful for gynecologists and scholars working in the field of human female reproductive health: bridging the gaps from genomics to phenomics, combining/merging the various specialties via multi-OMICS integration.

OMICS Discipline	Example	References
Genomics	Cancer genomics	[10,11]
	Nutrigenomics	[17,18]
	Epigenomics	[19]
	Nutriepigenomics	[20,21]
	Exosome genomics	[16]
Proteomics	Cancer proteomics	[12–15]
	Exosome proteomics	[16]
Transcriptomics	Cancer transcriptomics	[22]
Cytomics	Cancer cytomics	[22–24]
Metabolomics/metabonomics	Metabolomics	[16,35]
	Exometabolomics/microbial exometabolomics	
Microbiomics	Vaginal and maternal microbiomics; microbial culturomics	[25–33]
Pharmacomicrobiomics and pharmacoculturomics	Cancer pharmacomicrobiomics	[34]
Multi-omics	Multi-omics of preterm birth	[36]
	Multi-omics of gynecological cancers	[37]
Phenomics	Cancer and infertility phenomics	[38–44]

Int. J. Environ. Res. Public Health **2021**, *18*, 5058

6 of 13

4. Gynecology and Computational/Clinical Data

There exist several high-quality databases in the field of gynecology. For instance, the PREgnancy Care Integrating translational Science, Everywhere (PRECISE) database [45] is a unique resource for gynecologists and scholars working in the field of female reproductive health. This registry represents a core infrastructure of a data collection platform that has been developed and devised based on outcome data harmonization, and data recording and reporting standardization, relying on open-access technologies and clinical collaborating networks. It combines socio-demographic and clinical information with biological/biomedical parameters, including Human Immunodeficiency Virus (HIV) status, tests performed (including blood pressure measurement), therapeutical management received (for instance, iron therapy or uterotonic usage after delivery), postpartum maternal assessment within 48 h of birth, and newborn resuscitation, immediate skin-to-skin contact, and immediate drying, among others.

Other databases relevant to gynecology include: the Center of Excellence in Minimally Invasive Gynecology (COEMIG) registry, the Patient Centered Results for Uterine Fibroids (COMPARE-UF) database, the National Quality Registry for Gynaecological Surgery (GynOp) registry, the PanCare Studies in Fertility and Ototoxicity to Improve Quality of Life after Cancer during Childhood, Adolescence and Young Adulthood (PanCareLIFE) database, the Pelvic Floor Disorders (PFD) registry and the Society for Assisted Reproductive Technology (SART) registry.

Registries can also include patient-reported outcome measures (PROMs), thus providing fundamental information about a patient's perspective on gynecological management in terms of adherence and compliance to treatment and therapeutical preferences [46]. As such, they represent valuable resources for physicians and gynecologists. Some of the major big data-based databases available in the field are summarized in Table 3.

Table 3. An overview of major big data-based databases useful for gynecologists and scholars working in the field of human female reproductive health.

Database	Extended Title
COEMIG	Center of Excellence in Minimally Invasive Gynecology
COMPARE-UF	Patient Centered Results for Uterine Fibroids
GynOp	National Quality Registry for Gynaecological Surgery
PanCareLIFE	PanCare Studies in Fertility and Ototoxicity to Improve Quality of Life after Cancer during Childhood, Adolescence and Young Adulthood
PFD Registry	Pelvic Floor Disorders
PRECISE Registry	PREgnancy Care Integrating translational Science, Everywhere
SART Registry	Society for Assisted Reproductive Technology

5. Gynecology and Imaging/Wearable Sensor Data

A new big data-based specialty, termed radiomics, is emerging, which exploits radiographic medical images utilizing artificial intelligence-based algorithms, such as data-characterization or automatic segmentation and interpretation techniques, to uncover hidden trends and patterns and extract mineable information [47] in order to predict, for instance, the prognosis of patients suffering from malignancies [48]. For instance, Rizzo et al. [48] carried out a systematic review of the literature to evaluate the predictive power of radiomics in properly estimating overall survival and progression free survival in ovarian cancer patients. Based on six articles, statistically significant associations between radiomic features and survival rates could be found.

An unprecedented wealth of data can also be obtained by using wearable sensors and related technologies [49], which can collect information concerning women's periods, ovulation cycles, and breastfeeding habits, as well as other bodily functions. As such, IoT and IoMT can really advance gender equality and empower women. For instance, Runkle et al. [50] investigated the use of such sensors during prenatal care. Authors conducted

a web questionnaire-based study, utilizing a 21-item survey, which was delivered to 28 obstetrics and gynecology or family medicine providers. Another 21-item paper survey was administered to a sample of 103 pregnant women. Authors were able to find high acceptability levels of mobile and digital technologies both among healthcare providers and patients. Noteworthy, approximately 70% of women were in favor of utilizing smartphones for receiving individualized suggestions during their pregnancy, expressing their willingness to change their behaviors and adopt healthier lifestyles. Women were also interested in health and environmental tracking and monitoring. Similarly, healthcare providers expressed interest in mobile and smart technologies being used in tandem with 'classical' ones, stating that they could help them in their daily medical practice and feeling that they would have an increasing role in tele-gynecology.

In another study, Niela-Vilén et al. [51] were able to assess pregnant women's wellbeing during the ongoing coronavirus (COVID-19) pandemic by means of smartwatch technology. Authors performed a longitudinal cohort study recruiting two cohorts of women: one during the first COVID-19 wave with a history of preterm births or late miscarriages, and one during the second COVID-19 wave without pregnancy loss and with a history of full-term births. Physical activity practice significantly decreased, due to COVID-19 induced public health measures (such as quarantine, lockdown, and social distancing) and pregnancy, whereas sleep time was impacted by pregnancy but not by public health restrictions. Heart rate variability changed as well with respect to baseline values, with stress levels being significantly increased.

6. Gynecology and Infodemiology

'Infodemiology' (a portmanteau of 'information' and 'epidemiology') and 'infoveillance' (a combination of the words 'information' and 'surveillance') are the novel, emerging sciences of "distribution and determinants of information in an electronic medium, specifically the internet, or in a population, with the ultimate aim to inform and improve public health and public policy" [52].

Despite the increasing importance of the web in contemporary society and its usage by patients [53,54], in the existing scholarly literature, very few studies have addressed such a topic. For instance, Rezniczek et al. [55] retrieved and quantitatively assessed a set of 672 websites related to gynecological topics from several countries, including Switzerland (n = 49), Austria (n = 57), and Germany (n = 566). Websites were objectively ranked utilizing a composite panel of criteria: namely, the 2-item Google search ranking, a detailed, ad hoc devised checklist regarding technical aspects (in total, 11 items), web navigation and surfing (in total, 8 items), and website content (in total, 6 items) for an overall 26-point score. Scores were compared nationally (within a country) and internationally (between countries). The average score of websites was computed to be 13.8 ± 3.3, with only 4.2% of websites being rated as good. Most online resources (61.8%) were judged to be fair. Professional affiliation (academic versus non-academic), size and dimension of the institution (single versus consortium), and geographical location were found to be the determinants of the overall score. At the multivariate regression analysis, all sub-scores (Google search ranking, technical aspects, web navigation and surfing, and website content) could statistically predict the overall score, with the latter sub-score being the most significant one. Finally, all sub-scores exhibited statistically robust correlations among each other.

Han et al. [56] carried out a cross-sectional study of internet websites related to abortion, in order to investigate its relationship with psychological wellbeing and safety, depression, and mental health, as well as infertility. For this purpose, they exploited the Google AdWords tool to retrieve the keywords most associated with the topics under study and Google Trends to monitor and track Google's search engine to build a list of databases of websites commonly consulted. These sites were then categorized and assessed in terms of content and attitude towards abortion, reliability and accuracy, and type of online resource. Authors collected and analyzed 316 websites: more than half of them (57.3%) were unbiased and neutral towards abortion, with approximately one third (35.4%)

being clearly anti-choice, and only 7.3% being pro-choice. The average trustworthiness score varied among websites according to their content and attitude, being higher in neutral and pro-choice sites, with respect to anti-choice ones.

Endometriosis is a non-lethal but highly disabling gynecological condition characterized by infertility, pelvic pain and fatigue, allergies, and gastro-enterological problems. Hirsch et al. [57] mined five popular internet search engines (namely, aol.com, ask.com, bing.com, google.com, and yahoo.com), searching for endometriosis-related web resources. The authors were able to identify 54 relevant websites out of an initial list of 750 websites. These digital resources were assessed in terms of content credibility, quality, readability, and accuracy. The authors found that over a third and approximately half of these websites failed to attribute authorship and report the sources of information or to back up the online material with scholarly references, respectively. Noteworthy, authors could not identify any website disseminating endometriosis-related information written in plain English, with very few websites being judged as high-quality resources and with only one website displaying reliable, evidence-based information (namely, evidentlycochrane.net). Generally speaking, websites focused on diagnosis of endometriosis, without discussing other topics relevant for patients suffering from this gynecological disorder and mentioning content limitations. In conclusion, no website succeeded in achieving a high score in all four domains assessed by the authors (credibility, quality, readability, and accuracy).

Adawi et al. [58] conducted an infodemiological study in order to assess endometriosis-related digital activities in terms of web searches, after that an Italian celebrity, the television presenter Rossella Brescia, publicly announced that she was affected by this gynecological issue. More specifically, authors carried out a retrospective search of digital seeking behavior related to endometriosis in Italy, over the previous 5 years, exploiting Google Trends (section Google News) and utilizing "endometriosis" as the keyword and "search topic" as the search strategy option. The impact of media coverage and news dissemination on web behavior was also explored. The authors were able to find a statistically borderline significant association between Rossella Brescia and endometriosis related web searches. The self-disclosure of the celebrity impacted digital behavior, resulting in bursts of web activity.

In conclusion, based on this literature overview, we can state that physicians, and in particular gynecologists, should become aware of the increasing role of new media and social networks in contemporary society. Patients surf the web to retrieve information about their condition and share details with other patients, creating online communities [59,60]. However, patients have a higher chance of encountering websites spreading biased, inaccurate, and misleading material. As such, doctors should significantly strengthen their presence online and make better use of this means of interaction with patients [60,61]. Physicians' communication, while preserving trustworthiness and scholarly rigor, should be accessible to the lay public, aiming at improving their health literacy, promoting healthy lifestyles, and increasing and enhancing the awareness of preventative strategies. To this end, advocacy and the celebrity effect could be exploited [58,62].

7. An Overview of Use of Big Data in the Field of Gynecology

In summary, big data has so far been exploited in different sub-fields of gynecology, ranging from human reproductive health, physiology, and physiopathology (including pregnancy, infertility, and abortion) to assisted reproduction, sexually transmitted diseases (e.g., HPV), reproductive disorders (like endometriosis), and malignancies (endometrial, cervical, and ovarian cancers). Big data has enabled the collection of insightful information concerning early disease detection, stratification, and prognosis, as well as about PROMs. These are briefly summarized in Table 4.

Int. J. Environ. Res. Public Health **2021**, 18, 5058

9 of 13

Table 4. Aspects and sub-fields of gynecology that can benefit from the use of big data.

Gynecology Sub-Field Potentially Interested by Big Data	Examples
Reproductive health (pregnancy, infertility, abortion, endometriosis) and link to mental health and psychological well-being	Physiological and physio-pathological insights Health-related literacy
Assisted reproduction	Public interest and health-related literacy
Sexually transmitted diseases (such as HPV)	Health-related literacy
Gynecological cancers (endometrial, cervical, and ovarian cancers)	Molecular and cellular characterizations Clinical outcomes Health-related literacy

8. Pitfalls and Limitations of Big Data

Despite its strengths, big data is not immune from shortcomings and drawbacks [63]. These can be both technical and ethical/legal/bio-ethical. Among technical challenges, data portability refers to the possibility of transferring data from a given device, and format to other devices and formats, preserving their readability. Another particular technical challenge of Big Data is, given their variety and variability, to combine and organize them within a unique, coherent framework that can really advance our understanding of female reproductive health and gynecological disorders. This fragmentation and heterogeneity of data sources and types may challenge the classical concepts of ethical consent, data access, data sharing, and the distinction between data owner/legal custodian/buyer and data user (for example, researcher, stakeholder, citizen, etc.)—which is becoming more and more blurred. Data privacy (i.e., handling data in such a way as to minimize the risk of data theft/breach and to ensure the protection of sensitive information), data protection especially for unstructured data (personally identifiable), data integrity, data transparency, data replicability/reproducibility and other (bio-)ethical concerns represent further pitfalls of big data. In more detail, these include risk assessment and risk analysis, and uncertainty of communication to lay people, among others [64]. These are briefly summarized in Table 5.

Table 5. The main pitfalls and limitations plaguing the usage of big data.

Potential Pitfall
Technical Challenge
Data integration and combination Data portability
Ethical/Legal Challenge
Blurred distinction between data owner and data user Ethical consent Data privacy and data protection Data portability Data sharing Data integrity Data transparency Data replicability/reproducibility

Specifically concerning gynecology, and regarding human reproductive health in general, information related to a patient's health status is particularly sensitive. Pregnancy is also a period in which women could experience privacy concerns. However, some studies [50] have demonstrated high levels of acceptability among patients towards big data and computational/digital technologies (such as mobile sensors), with pregnant women being comfortable with utilizing such devices and sharing information from these tools with their physicians [50].

However, few data concerning these attitudes are available and there is a gap in the literature concerning this specific topic. Future high-quality studies should be conducted examining the determinants of acceptability of big data based smart technologies among women.

9. Areas for Future Research

While some big data-based resources, such as large, nation-wide epidemiological surveys and claim databases, are already available for gynecologists, they are not adequately trained to exploit such tools. Medical students, residents, and future practitioners should be instructed and educated about the opportunities offered by big and smart data. This is of paramount importance, in that data mining can potentially result in a high number of parameters being statistically significant and gynecologists should assist in the process of decision-making and critical and meaningful interpretation of results, being able to distinguish between statistical and clinical significance. Involving practitioners in translational research could facilitate and expedite the transition from bench to bedside, helping to fill in several gaps of knowledge, conjugating, integrating and harmonizing discovery bioscience and clinical care.

10. Conclusions

Within the new theoretical framework of the 'disruptive innovation medical era', big and smart data are anticipated to revolutionize different medical specializations, including gynecology, shedding light on female reproductive health, both in terms of physiology and pathophysiology. More specifically, big data appears to be promising, in the field of gynecology, to potentially increase its accuracy and precision, stratify patients, provide opportunities for personalized treatment options rather than delivering a package of "one-size-fits-it-all" healthcare management provisions, and enhance its effectiveness at each stage (health promotion, prevention, diagnosis, prognosis, and therapeutics). However, research to overcome the technical and ethical limitations and pitfalls of big data is warranted.

References

1. Bragazzi, N.L. From P0 to P6 medicine, a model of highly participatory, narrative, interactive, and "augmented" medicine: Some considerations on Salvatore Iaconesi's clinical story. *Patient Prefer. Adherence* **2013**, *7*, 353–359. [CrossRef] [PubMed]
2. Wiweko, B.; Zakirah, S.C. Future of Health Services: The Role of Physicians in the Disruptive Era. *J. Hum. Reprod. Sci.* **2020**, *13*, 250–256. [CrossRef] [PubMed]
3. Oh, H.; Rizo, C.; Enkin, M.; Jadad, A. What is eHealth (3): A systematic review of published definitions. *J. Med. Internet Res.* **2005**, *7*, e1. [CrossRef] [PubMed]
4. Singh, S.; Singh, A. Current status of nanomedicine and nanosurgery. *Anesth. Essays Res.* **2013**, *7*, 237–242. [CrossRef] [PubMed]
5. Hou, C.; Zheng, J.; Li, Z.; Qi, X.; Tian, Y.; Zhang, M.; Zhang, J.; Huang, X. Printing 3D vagina tissue analogues with vagina decellularized extracellular matrix bioink. *Int. J. Biol. Macromol.* **2021**, *180*, 177–186. [CrossRef]
6. Dini, G.; Bragazzi, N.L.; Montecucco, A.; Toletone, A.; Debarbieri, N.; Durando, P. Big Data in occupational medicine: The convergence of -omics sciences, participatory research and e-health. *Med. Lav.* **2019**, *110*, 102–114.
7. Krumholz, H.M. Big data and new knowledge in medicine: The thinking, training, and tools needed for a learning health system. *Health Aff.* **2014**, *33*, 1163–1170. [CrossRef]

Int. J. Environ. Res. Public Health **2021**, *18*, 5058

11 of 13

8. Delude, C.M. Deep phenotyping: The details of disease. *Nature* **2015**, *527*, S14–S15. [CrossRef] [PubMed]

9. Robinson, P.N. Deep phenotyping for precision medicine. *Hum. Mutat.* **2012**, *33*, 777–780. [CrossRef]

10. Muinao, T.; Deka Boruah, H.P.; Pal, M. Multi-biomarker panel signature as the key to diagnosis of ovarian cancer. *Heliyon* **2019**, *5*, e02826. [CrossRef]

11. Talhouk, A.; George, J.; Wang, C.; Budden, T.; Tan, T.Z.; Chiu, D.S.; Kommoss, S.; San Leong, H.; Chen, S.; Intermaggio, M.P.; et al. Development and Validation of the Gene Expression Predictor of High-grade Serous Ovarian Carcinoma Molecular SubTYPE (PrOTYPE). *Clin. Cancer Res.* **2020**, *26*, 5411–5423. [CrossRef] [PubMed]

12. Ahn, H.S.; Yeom, J.; Yu, J.; Kwon, Y.I.; Kim, J.H.; Kim, K. Convergence of Plasma Metabolomics and Proteomics Analysis to Discover Signatures of High-Grade Serous Ovarian Cancer. *Cancers* **2020**, *12*, 3447. [CrossRef] [PubMed]

13. Daly, M.B.; Ozols, R.F. The search for predictive patterns in ovarian cancer: Proteomics meets bioinformatics. *Cancer Cell* **2002**, *1*, 111–112. [CrossRef]

14. Lekhwani, S.; Shankar, V.; Vaswani, N.D. Proteomics in obstetrics and gynecology. *Indian J. Hum. Genet.* **2011**, *17*, 3–6. [CrossRef]

15. Timms, J.F.; Arslan-Low, E.; Kabir, M.; Worthington, J.; Camuzeaux, S.; Sinclair, J.; Szaub, J.; Afrough, B.; Podust, V.N.; Fourkala, E.O.; et al. Discovery of serum biomarkers of ovarian cancer using complementary proteomic profiling strategies. *Proteom. Clin. Appl.* **2014**, *8*, 982–993. [CrossRef] [PubMed]

16. Esfandyari, S.; Elkafas, H.; Chugh, R.M.; Park, H.S.; Navarro, A.; Al-Hendy, A. Exosomes as Biomarkers for Female Reproductive Diseases Diagnosis and Therapy. *Int. J. Mol. Sci.* **2021**, *22*, 2165. [CrossRef]

17. Bulloch, R.E.; Wall, C.R.; McCowan, L.M.E.; Taylor, R.S.; Roberts, C.T.; Thompson, J.M.D. The Effect of Interactions between Folic Acid Supplementation and One Carbon Metabolism Gene Variants on Small-for-Gestational-Age Births in the Screening for Pregnancy Endpoints (SCOPE) Cohort Study. *Nutrients* **2020**, *12*, 1677. [CrossRef]

18. Sellami, M.; Bragazzi, N.L. Nutrigenomics and Breast Cancer: State-of-Art, Future Perspectives and Insights for Prevention. *Nutrients* **2020**, *12*, 512. [CrossRef]

19. Ronkainen, J.; Heiskala, A.; Vehmeijer, F.O.L.; Lowry, E.; Caramaschi, D.; Estrada Gutierrez, G.; Heiss, J.A.; Hummel, N.; Keikkala, E.; Kvist, T.; et al. Maternal haemoglobin levels in pregnancy and child DNA methylation: A study in the pregnancy and childhood epigenetics consortium. *Epigenetics* **2021**, 1–13. [CrossRef]

20. Do, W.L.; Whitsel, E.A.; Costeira, R.; Masachs, O.M.; Le Roy, C.I.; Bell, J.T.R.; Staimez, L.; Stein, A.D.; Smith, A.K.; Horvath, S.; et al. Epigenome-wide association study of diet quality in the Women's Health Initiative and TwinsUK cohort. *Int. J. Epidemiol.* **2020**. [CrossRef]

21. Bordoni, L.; Gabbianelli, R. Primers on nutrigenetics and nutri(epi)genomics: Origins and development of precision nutrition. *Biochimie* **2019**, *160*, 156–171. [CrossRef]

22. Veenstra, T.D. Omics in Systems Biology: Current Progress and Future Outlook. *Proteomics* **2021**, *21*, e2000235. [CrossRef]

23. Allain, G. Genomics, proteomics, cytomics, and genetic. *MLO. Med. Lab. Obs.* **2012**, *44*, 36–37.

24. Ke, J.; Shen, Y.; Lu, Y.; Deng, J.; Wright, J.D.; Zhang, Y.; Huang, Q.; Wang, D.; Jing, N.; Liang, X.; et al. Quantitative analysis of abnormalities in gynecologic cytopathology with deep learning. *Lab. Investig.* **2021**. [CrossRef]

25. Tamarelle, J.; Thiébaut, A.C.M.; de Barbeyrac, B.; Bébéar, C.; Ravel, J.; Delarocque-Astagneau, E. The vaginal microbiota and its association with human papillomavirus, Chlamydia trachomatis, Neisseria gonorrhoeae and Mycoplasma genitalium infections: A systematic review and meta-analysis. *Clin. Microbiol. Infect.* **2019**, *25*, 35–47. [CrossRef] [PubMed]

26. Singh, T.; Neal, A.S.; Moatamed, N.A.; Memarzadeh, S. Exploring the Potential of Drug Response Assays for Precision Medicine in Ovarian Cancer. *Int. J. Mol. Sci.* **2020**, *22*, 305. [CrossRef]

27. Mancabelli, L.; Tarracchini, C.; Milani, C.; Lugli, G.A.; Fontana, F.; Turroni, F.; van Sinderen, D.; Ventura, M. Vaginotypes of the human vaginal microbiome. *Environ. Microbiol.* **2021**. [CrossRef]

28. Liu, J.; Luo, M.; Zhang, Y.; Cao, G.; Wang, S. Association of high-risk human papillomavirus infection duration and cervical lesions with vaginal microbiota composition. *Ann. Transl. Med.* **2020**, *8*, 1161. [CrossRef]

29. Lee, C.Y.; Cheu, R.K.; Lemke, M.M.; Gustin, A.T.; France, M.T.; Hampel, B.; Thurman, A.R.; Doncel, G.F.; Ravel, J.; Klatt, N.R.; et al. Quantitative modeling predicts mechanistic links between pre-treatment microbiome composition and metronidazole efficacy in bacterial vaginosis. *Nat. Commun.* **2020**, *11*, 6147. [CrossRef]

30. Di Simone, N.; Santamaria Ortiz, A.; Specchia, M.; Tersigni, C.; Villa, P.; Gasbarrini, A.; Scambia, G.; D'Ippolito, S. Recent Insights on the Maternal Microbiota: Impact on Pregnancy Outcomes. *Front. Immunol.* **2020**, *11*, 528202. [CrossRef]

31. Fettweis, J.M.; Serrano, M.G.; Brooks, J.P.; Edwards, D.J.; Girerd, P.H.; Parikh, H.I.; Huang, B.; Arodz, T.J.; Edupuganti, L.; Glascock, A.L.; et al. The vaginal microbiome and preterm birth. *Nat. Med.* **2019**, *25*, 1012–1021. [CrossRef]

32. Susic, D.; Davis, G.; O' Sullivan, A.J.; McGovern, E.; Harris, K.; Roberts, L.M.; Craig, M.E.; Mangos, G.; Hold, G.L.; El-Omar, E.M.; et al. Microbiome Understanding in Maternity Study (MUMS), an Australian prospective longitudinal cohort study of maternal and infant microbiota: Study protocol. *BMJ Open* **2020**, *10*, e040189.

33. Diop, K.; Diop, A.; Levasseur, A.; Mediannikov, O.; Robert, C.; Armstrong, N.; Couderc, C.; Bretelle, F.; Raoult, D.; Fournier, P.E.; et al. Microbial Culturomics Broadens Human Vaginal Flora Diversity: Genome Sequence and Description of Prevotella lascolaii sp. nov. Isolated from a Patient with Bacterial Vaginosis. *OMICS* **2018**, *22*, 210–222. [CrossRef]

34. ElRakaiby, M.; Dutilh, B.E.; Rizkallah, M.R.; Boleij, A.; Cole, J.N.; Aziz, R.K. Pharmacomicrobiomics: The impact of human microbiome variations on systems pharmacology and personalized therapeutics. *OMICS* **2014**, *18*, 402–414. [CrossRef]

35. Harden, S.L.; Zhou, J.; Gharanei, S.; Diniz-da-Costa, M.; Lucas, E.S.; Cui, L.; Murakami, K.; Fang, J.; Chen, Q.; Brosens, J.J.; et al. Exometabolomic Analysis of Decidualizing Human Endometrial Stromal and Perivascular Cells. *Front. Cell. Dev. Biol.* **2021**. [CrossRef]

36. Jehan, F.; Sazawal, S.; Baqui, A.H.; Nisar, M.I.; Dhingra, U.; Khanam, R.; Ilyas, M.; Dutta, A.; Mitra, D.K.; Mehmood, U.; et al. Multiomics Characterization of Preterm Birth in Low- and Middle-Income Countries. *JAMA. Netw. Open* **2020**, *3*, e2029655. [CrossRef]

37. Li, J.Y.; Li, C.J.; Lin, L.T.; Tsui, K.H. Multi-Omics Analysis Identifying Key Biomarkers in Ovarian Cancer. *Cancer Control.* **2020**, *27*, 1073274820976671. [CrossRef]

38. Westergaard, D.; Nielsen, A.P.; Mortensen, L.H.; Nielsen, H.S.; Brunak, S. Phenome-Wide Analysis of Short- and Long-Run Disease Incidence Following Recurrent Pregnancy Loss Using Data From a 39-Year Period. *J. Am. Heart Assoc.* **2020**, *9*, e015069. [CrossRef]

39. Joo, Y.Y.; Actkins, K.; Pacheco, J.A.; Basile, A.O.; Carroll, R.; Crosslin, D.R.; Hakonarson, H.; Harley, J.B. A Polygenic and Phenotypic Risk Prediction for Polycystic Ovary Syndrome Evaluated by Phenome-Wide Association Studies. *J. Clin. Endocrinol. Metab.* **2020**, *105*, 1918–1936. [CrossRef] [PubMed]

40. Vitonis, A.F.; Vincent, K.; Rahmioglu, N.; Fassbender, A.; Buck Louis, G.M.; Hummelshoj, L.; Giudice, L.C.; Stratton, P.; Adamson, G.D.; Becker, C.M.; et al. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonization Project: II. Clinical and covariate phenotype data collection in endometriosis research. *Fertil. Steril.* **2014**, *102*, 1223–1232. [CrossRef]

41. Rahmioglu, N.; Fassbender, A.; Vitonis, A.F.; Tworoger, S.S.; Hummelshoj, L.; D'Hooghe, T.M.; Adamson, G.D.; Giudice, L.C.; Becker, C.M.; Zondervan, K.T.; et al. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonization Project: III. Fluid biospecimen collection, processing, and storage in endometriosis research. *Fertil. Steril.* **2014**, *102*, 1233–1243. [CrossRef] [PubMed]

42. Fassbender, A.; Rahmioglu, N.; Vitonis, A.F.; Viganò, P.; Giudice, L.C.; D'Hooghe, T.M.; Hummelshoj, L.; Adamson, G.D.; Becker, C.M.; Missmer, S.A.; et al. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonisation Project: IV. Tissue collection, processing, and storage in endometriosis research. *Fertil. Steril.* **2014**, *102*, 1244–1253. [CrossRef]

43. Becker, C.M.; Laufer, M.R.; Stratton, P.; Hummelshoj, L.; Missmer, S.A.; Zondervan, K.T.; Adamson, G.D.; WERF EPHect Working Group. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonisation Project: I. Surgical phenotype data collection in endometriosis research. *Fertil. Steril.* **2014**, *102*, 1213–1222. [CrossRef]

44. Viganò, P.; Somigliana, E.; Panina, P.; Rabellotti, E.; Vercellini, P.; Candiani, M. Principles of phenomics in endometriosis. *Hum. Reprod. Update* **2012**, *18*, 248–259. [CrossRef]

45. Magee, L.A.; Strang, A.; Li, L.; Tu, D.; Tumtaweetikul, W.; Craik, R.; Daniele, M.; Etyang, A.K.; D'Alessandro, U.; Ogochukwu, O.; et al. The PRECISE (PREgnancy Care Integrating translational Science, Everywhere) database: Open-access data collection in maternal and newborn health. *Reprod. Health* **2020**, *17*, 50. [CrossRef]

46. Erekson, E.A.; Iglesia, C.B. Improving Patient Outcomes in Gynecology: The Role of Large Data Registries and Big Data Analytics. *J. Minim. Invasive Gynecol.* **2015**, *22*, 1124–1129. [CrossRef] [PubMed]

47. Rizzo, S.; Botta, F.; Raimondi, S.; Origgi, D.; Fanciullo, C.; Morganti, A.G.; Bellomi, M. Radiomics: The facts and the challenges of image analysis. *Eur. Radiol. Exp.* **2018**, *2*, 36. [CrossRef]

48. Rizzo, S.; Manganaro, L.; Dolciami, M.; Gasparri, M.L.; Papadia, A.; Del Grande, F. Computed Tomography Based Radiomics as a Predictor of Survival in Ovarian Cancer Patients: A Systematic Review. *Cancers* **2021**, *13*, 573. [CrossRef]

49. Yoldemir, T. Internet of Things and women's health. *Climacteric* **2020**, *23*, 423–425. [CrossRef] [PubMed]

50. Runkle, J.; Sugg, M.; Boase, D.; Galvin, S.L.C.; Coulson, C. Use of wearable sensors for pregnancy health and environmental monitoring: Descriptive findings from the perspective of patients and providers. *Digit. Health* **2019**, *5*, 2055207619828220. [CrossRef]

51. Niela-Vilén, H.; Auxier, J.; Ekholm, E.; Sarhaddi, F.; Asgari Mehrabadi, M.; Mahmoudzadeh, A.; Azimi, I.; Liljeberg, P.; Rahmani, A.M.; Axelin, A. Pregnant women's daily patterns of well-being before and during the COVID-19 pandemic in Finland: Longitudinal monitoring through smartwatch technology. *PLoS ONE* **2021**, *16*, e0246494. [CrossRef] [PubMed]

52. Eysenbach, G. Infodemiology and infoveillance: Framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J. Med. Internet Res.* **2009**, *11*, e11. [CrossRef]

53. Suchard, M.A.; Hadfield, R.; Elliott, T.; Kennedy, S. Beyond providing information: The Internet as a research tool in reproductive medicine. *Hum. Reprod.* **1998**, *13*, 6–7. [CrossRef] [PubMed]

54. Deevey, S. Endometriosis: Internet resources. *Med. Ref. Serv. Q.* **2005**, *24*, 67–76. [CrossRef]

55. Rezniczek, G.A.; Küppers, L.; Heuer, H.; Hefler, L.A.; Buerkle, B.; Tempfer, C.B. Quality of websites of obstetrics and gynecology departments: A cross-sectional study. *BMC Pregnancy Childbirth* **2015**, *15*, 103. [CrossRef] [PubMed]

56. Han, L.; Boniface, E.R.; Han, L.Y.; Albright, J.; Doty, N.; Darney, B.G. The Abortion Web Ecosystem: Cross-Sectional Analysis of Trustworthiness and Bias. *J. Med. Internet Res.* **2020**, *22*, e20619. [CrossRef] [PubMed]

57. Hirsch, M.; Aggarwal, S.; Barker, C.; Davis, C.J.; Duffy, J.M.N. Googling endometriosis: A systematic review of information available on the Internet. *Am. J. Obstet. Gynecol.* **2017**, *216*, 451–458.e1. [CrossRef]

58. Adawi, M.; Martini, M.; Watad, A.; Tornali, C.; Barberis, I.; Borromeo, F.; Bragazzi, N.L. Celebrity effect in public health and bioethics: An "Italian case" in endometriosis prevention communication. *Acta Med. Mediterr.* **2019**, *35*, 913.

Int. J. Environ. Res. Public Health **2021**, *18*, 5058

13 of 13

59. Wilson, S.; Mogan, S.; Kaur, K. Understanding the role of Facebook to support women with endometriosis: A Malaysian perspective. *Int. J. Nurs. Pract.* **2020**, *26*, e12833. [CrossRef]

60. Carneiro, M.M.; Farace, B.L.; Ribeiro, L.S.C.; Silverio, R.C.C.C.; Moreira, T.; Filho, A.L.D.S.; Baroni, A.L.L.R.; Ferreira, M.C.F. Using social media to educate women and healthcare providers on endometriosis: Preliminary results. *JBRA. Assist. Reprod.* **2020**, *24*, 9–12. [CrossRef]

61. Shadbolt, N.A.; Parker, M.A.; Orthia, L.A. Communicating endometriosis with young women to decrease diagnosis time. *Health Promot. J. Aust.* **2013**, *24*, 151–154. [CrossRef]

62. Blakemore, J.K.; Bayer, A.H.; Smith, M.B.; Grifo, J.A. Infertility influencers: An analysis of information and influence in the fertility webspace. *J. Assist. Reprod. Genet.* **2020**, *37*, 1371–1378. [CrossRef] [PubMed]

63. Ebbesen, M.; Jensen, T.G. Nanomedicine: Techniques, potentials, and ethical implications. *J. Biomed. Biotechnol.* **2006**, *2006*, 51516. [CrossRef]

64. Nateghi, R.; Aven, T. Risk Analysis in the Age of Big Data: The Promises and Pitfalls. *Risk Anal.* **2021**. [CrossRef] [PubMed]