Abstract architectural framework and methodology, underlying benefits, describes

AbstractTheprocess of examining huge and different data sets is known as Big dataanalytics. It is used to uncover hidden patterns correlation and other insightswith today’s technology. Big data is mainly used to analyse data and get answerfor the analysed data immediately. A database isa collection of information that is organized sothat it can be easily accessed, managed and updated. A relational database, more restrictively, is a collection of schemas, tables, queries, reports, views, and other elements. The outcome of the research paper is how bigdata analytics and database is applicable in medical sector on Promises . This paper provides a broad overview of big data analytics forhealthcare researchers and practitioners and health database organization.

Keyword: Big data analytics, healthcare, Database. I. INTRODUCTIONThehealthcare industry historically has generated large amounts of data, driven byrecord keeping, compliance & regulatory requirements, and patient care. While most data is stored in hard copy form, the current trend is towards rapiddigitization of these large amounts of data. 8 Healthcareorganizations are now presented with a flood of data. Everyone in theorganization—from the CEO to the individual clinician—needs a way to turn allof that raw information into targeted, actionable knowledge. Making sense ofthe raw data on its own, without tools and processes to guide the process, canbe overwhelming.

Here, weare going to compare the “ Big data and Data Base” in healthcare. II. LITERATURE REVIEW This paper includes thefundamental field of big data analytics in healthcare, outlines an architectural framework andmethodology, underlying benefits, describes examples reported in the

literature, briefly explains the challenges and offers conclusion. This paper providesthe overview of big data analytics for healthcare research and practitioners. 2WullianallurRaghupathi and Viju Raghupathi (2014). HasAdvantagesof healthcare: There are variousmerits of using big data analytics in healthcare. By digitizing, combining andeffectively using big data, healthcare organizations ranging from asingle-physician offices to large hospital networks benefited through big dataanalytics. McKinsey believes big data could help reduce waste and inefficiencyin the following three areas:· Clinical operation· Research anddevelopment· Public Health Ø Evidence basedmedicineØ Genomic AnalyticsØ Pre-AdjudicationFraud AnalysisØ Device or RemoteMonitoringØ Patient ProfileAnalyticsWullianallurRaghupathi and Viju Raghupathi (2014).

He has also given explanation using 4″vs" of big data analytics in healthcare. The4 " vs" areVolume, Velocity and Variety are the primarycharacteristics. Veracity is the goalØ ArchitectureFramework WullianallurRaghupathi and Viju Raghupathi (2014). He has also given explanation using architecture framework of big dataanalytics in healthcare. Fig: 1Architectural FrameworkØ Outline of big data analytics inhealthcare methodology Wullianallur Raghupathi and VijuRaghupathi (2014).

He has also given explain the Outline of big data analyticsin healthcare methodology in the table formStep1 Concept statement · Establish need for big data analyticsproject in healthcare based on the " 4Vs". Step2 Proposal · What is the problem being addressed?· Why is it important and interesting?· Why big data analytics approach?

· Background material Step3 Methodology · Propositions · Variable selection · Data collection· ETL and data transformation · Platform/tool selection· Conceptual modelIII. COMPARATIVE STUDYØ BIG DATA Whatexactly is big data? A report delivered to the U. S. Congress in August 2012defines big data as " large volumes of high velocity, complex, and variable datathat require advanced techniques and technologies to enable the capture, storage, distribution, management and analysis of the information. 6Bigdata analytics in healthcare. Big data in healthcare is massive notonly because of its capacity but also because of the diversity of data typesand the speed at which it must be managed. By definition, big data inhealthcare refers to electronic health data sets are huge and difficult tomanage with traditional software and or hardware; nor can they be easilymanaged with traditional or common data management tools and methods.

Thehealthcare industry historically has generated large amounts of data, driven byrecord keeping, compliance & regulatory requirements, and patient care. While most data is stored in hard copy form, the current trend is toward rapiddigitization of these large amounts of data. Driven by mandatory requirementsand the potential to improve the quality of healthcare delivery meanwhilereducing the costs, these massive quantities of data (known as ' big data') holdthe promise of supporting a wide range of medical and healthcare functions, including among others clinical decision support, disease surveillance, andpopulation health management . Big dataencompasses such characteristics as variety, velocity and, with respectspecifically to healthcare, veracity. 7Advantagesto

healthcare By digitizing, combiningand effectively using big data, healthcare organizations ranging fromsingle-physician offices and multi-provider groups to large hospital networksand accountable care organizations stand to realize significant benefits.

Potential benefits include detecting diseases at earlier stages when they canbe treated more easily and effectively; managing specific individual andpopulation health and detecting health care fraud more quickly and efficiently. 9Numerous questions can be addressed with big data analytics. Certaindevelopments or outcomes may be predicted and/or estimated based on vastamounts of historical data, such as length of stay (LOS); patients who willchoose elective surgery; patients who likely will not benefit from surgery; complications; patients at risk for medical complications; patients at risk for sepsis, MRSA, C. difficile, or other hospital-acquired illness; illness/disease progression; patients at risk for advancement in disease states; causal factors ofillness/disease progression; and possible co morbid conditions (EMCConsulting).

Challenges At minimum, a big data analytics platform inhealthcare must support the key functions necessary for processing the data. The criteria for platform evaluation may include availability, continuity, easeof use, scalability, ability to manipulate at different levels of granularity, privacy and security enablement, and quality assurance. To succeed, big dataanalytics in healthcare needs to be packaged so it is menu driven, user-friendly and transparent. Real-time big data analytics is a keyrequirement in healthcare. The lag between data collection and processing hasto be addressed.

The dynamic availability of numerous analytics algorithms, models and methods in a pull-down type of menu is also necessary forlarge-scale adoption. The important managerial issues of ownership, governanceand standards have to be considered. And woven through these issues are thoseof continuous data acquisition and data cleansing. Health care data is rarelystandardized, often fragmented, or generated in legacy IT systems withincompatible formats .

This great challenge needs to be addressed as well. Ø DATABASETheterm database embracesmany different concepts: from paper records maintained by a single practitionerto the vast computerized collections of insurance claims for Medicarebeneficiaries. A database is a large collection of data in a computer, organized so that it can be expanded, updated, and retrieved rapidly forvarious uses. 10Key Attributes of DatabasesIn reviewing theconsiderable variation in databases that might be accessed, controlled, oracquired by HDOs, the committee sought a simple way to characterize them by keyattributes. It decided on two critical dimensions of databases: comprehensiveness and inclusiveness. 1 ComprehensivenessComprehensiveness refers to the amount ofinformation one has on an individual both for each patient encounter with thehealth care system and for all of a patient's encounters over time (USDHHS, 1991, refers to this as completeness). A record that is comprehensive contains: Demographic data, Administrative data, Health risks and Health status, Patientmedical history, Current management of health conditions, and Outcomes data. Inclusiveness Inclusiveness refers to which populations in a geographic area areincluded in a database.

Databases that aim to provide information on the healthof the community ought to include an enumeration of all residents of thecommunity (e. g., metropolitan area, state) 5The Concept of HDOsThe committee chose thephrase Health Database Organization (HDO) to refer to entitiesthat have access to databases and that have as their chief mission the publicrelease of data and of results of analyses done on the databases under theircontrol. For purposes of this report, prototypical HDOs have the characteristics outlinedin these properties · They operate under a single, common authority.

· They acquire and maintain information from a widevariety of sources in the health sector· Files accessible to HDOs willinclude person-identified orperson-identifiable data.· HDOs will serve a specific geographic area thatis defined chiefly by geographic or political boundaries (e. g., metropolitanarea, county, state) and will include those who reside in or receive servicesin that area, or both.· HDOs will process, store, analyze, and otherwise manipulatedata electronically. The Benefits of Health DatabaseThe gains expected from imaginativebut responsible uses of the information held by HDOs accrue not only to variousinterest groups but also to populations generally, whether in a metropolitan orsub state region, a given state, or the nation as a whole.

The size of thepotential benefits, whether to the community at large or to specific users, islikely to be a function of the comprehensivenessand inclusiveness of the databases the more comprehensive or inclusive the more powerful theinformation will be at every level and for every potential user and use. Broad-based BenefitsThe intent of many database and

HDOefforts today is to give regions a way to monitor and improve the value oftheir health care services and the well being of their residents. HDOs mightachieve this by making available information on access to care, costs, appropriateness, effectiveness, and quality of health care services andproviders · Access · Costs · Quality of care Delivery of health services Disease incidence and public health · Healthplanning IV. METHODOLOGYØ Bigdata analytics in healthcareThearticle " Big data analytics in healthcare: promise and potential written byWullianallur Raghupathi and Viju Raghupathi" uses the following methodology intheir research.

The below paragraphs displays the main stages of themethodology. 4
In Step1, the interdisciplinary bigdata analytics in healthcare team develops a ' concept statement'. The conceptstatement is followed by a description of the project's significance. Thehealthcare organization will note that there are trade-offs in terms ofalternative options, cost, scalability, etc. InStep2, the proposal development stage.

Here, more details are filled in. Based on the concept statement, several questions are addressed: What problemis being addressed? Why is it important and interesting to the healthcareprovider? What is the case for a ' big data' analytics approach? (Because thecomplexity and cost of big data analytics are significantly higher compared totraditional analytics approaches, it is important to justify their use). Theproject team also should provide background information on the problem domainas well as prior projects and research done in this domain. In Step 3, the steps in the methodology are fleshed out andimplemented. The concept statement is

broken down into a series ofpropositions. Simultaneously, the independent and dependent variables orindicators are identified. The data sources, as outlined in, are alsoidentified; the data is collected, described, and transformed in preparationfor analytics.

A very important step atthis point is platform/tool evaluation and selection. There are several optionsavailable, as indicated previously, including AWS Hadoop, Cloudera, and IBM BigInsights. The next step is to apply the various big data analytics techniquesto the data. This process differs from routine analytics only in that thetechniques are scaled up to large data sets. Through a series of iterations andwhat-if analyses, insight is gained from the big data analytics. From theinsight, informed decisions can be made.

In Step 4, the models and their findings are tested andvalidated and presented to stakeholders for action. Implementation is a stagedapproach with feedback loops built in ateach stage to minimize risk of failure. Ø Databasein healthcareEuro statStatistics explained what methodology is used in collecting data using databasein healthcare. The following are the methodologies3:· Householdbudget surveys- The Household budgetsurvey, abbreviated as HBS, is a national survey focusing on household’sexpenditure on goods and services· Administrativesources- The administrative source isthe register of units and data associated with an administrative regulation (or group of regulations), viewedas a source of statistical data.· Data collected for the purpose of national accounts- national accounts meansfocusing on the structure and evolution ofeconomies. · Data informationsystems available in health (and other) ministries / departments as well asother agencies involved in health care. V.

CONCLUSIONBig data analytics in healthcare providessophisticated technologies to gain knowledge about the clinical information. The healthcare industries are facing many challenges such as privacy, safeguarding security, standard of governance in providing secured dataanalytics in the area of all-over healthcare. Healthcare organizations andhealthcare industries are making steady progress in providing technologicalbased information in more transparent and as well as in an organized way. Health database organizations diagnose the health of the public; improve thequality of care in hospitals, clinics and various other care settings. Databasein clinical care should be maintained in a complete and accurate manner.

Anyinaccurate, missing data or out-of-date data are found it may cause a big harmto the overall sector. HDO’s can also contribute to improvements in quality ofcare by making information available to institutions and group of practitionersfor their use in quality assurance and quality improvement(QA/QI)programs forregional health planning. Hence we conclude by saying that big data analyticshas more storage than health database and it also reduces healthcare costs andit is more effective than the database in healthcare.