



White Paper

Integrated Data Analytics in an Era of Reform: Ten Observations From the Trenches

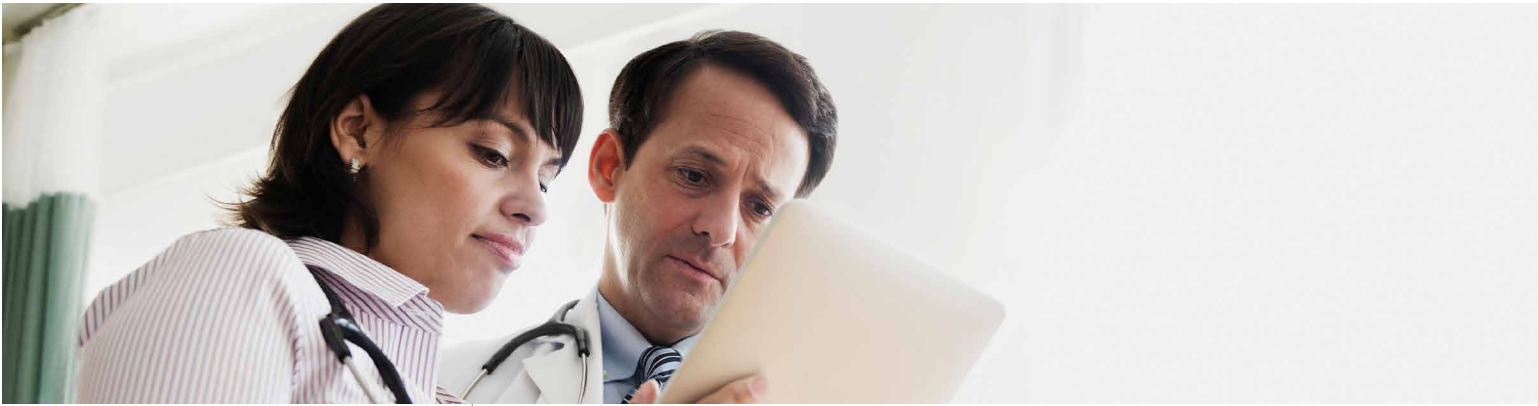
Grant Hoffman, MBA
Rebecca Molesworth, MHSA

November 2012



Table of Contents

- Executive Summary 1
- Driving Analytic Value With Integrated Data Streams
and Predictive Models 2
- Ten Observations From the Trenches 4
- Moving Forward 9
- About the Authors10



Executive Summary

Much has been recently written on the current regulatory, economic, and clinical climate in which health systems are delivering care in the United States. The potential of data transparency, patient empowerment, and quality improvement is paired with newfound intense market competition, complicated payer negotiations, eroding physician satisfaction, and evolving Centers for Medicare & Medicaid Services (CMS) administrative and payment mechanisms.

Analytics has become a key tool in helping providers manage their operations in this increasingly complex world. Health systems are seeking informational capabilities that preserve margins while still providing high-quality care to high-risk patient populations. This need for care delivery analytics has garnered federal attention. In the Institute of Medicine's report, "Better Care at Lower Cost,"¹ the researchers' first recommendation was that "data generated in the course of care delivery should be digitally collected, compiled, and protected as a reliable and accessible resource for care management, process improvement, public health, and the generation of new knowledge."

Using previously siloed data sets to closely monitor cost, quality, and risk via analytics is now the new normal for business. The following table summarizes the four major pressure sources driving organizations to seek integrated analytic solutions for decision support.

¹ *Institute of Medicine Report, Fall 2012*

The balance of these initiatives in each organization provides a unique set of circumstances and necessary analytical needs.

eHI/CHIME Survey Results

1. Lack of standardized data across systems (53.5%)
2. Lack of system infrastructure to support analytics (45.1%)
3. Cost of analytic software (43.7%)
4. Lack of data integration (38%)
5. Concerns about privacy/security of data (36.6%)
6. Physician alignment (29.6%)
7. Limited utility of the results to the organization (12.7%)

About the survey: 90+ responses were received from CIO members of CHIME during a four-week period. Of those, 36 percent were community hospitals, 30 percent were multi-hospital, 16 percent were academic medical centers, 8 percent were hospital/clinics, and 11 percent were categorized as “other.”

Examples of Health System Pressures Driving the Need for Analytic Solutions

Technical	Financial
<ul style="list-style-type: none"> ▪ ICD-10 implementation ▪ Meaningful Use reporting and attestation ▪ Integrating systems across facilities and acquisitions 	<ul style="list-style-type: none"> ▪ Rising costs/lower margins ▪ Taking on performance risk ▪ Lower reimbursements ▪ Federal penalties ▪ Referral leakage
Administrative	Clinical
<ul style="list-style-type: none"> ▪ Coordinating/shaping physician behaviors in a dispersed network ▪ Gaining market share ▪ Obtaining market intelligence on competition ▪ Tracking utilization and optimizing operations across multiple stakeholders 	<ul style="list-style-type: none"> ▪ Quality measurement and improvement ▪ Workflow management and improvement ▪ Access to and tracking use of evidence-based practices

The challenge for each organization is to understand how these pressures affect their organization and where to prioritize their scarce resources to drive the best ROI.

Driving Analytic Value With Integrated Data Streams and Predictive Models

Recent initiatives to push the health industry forward in its utilization of data collection, exchange, and analytics have flourished since the Health Information Technology for Economic and Clinical Health Act (HITECH) was enacted in 2009 as part of the American Recovery and Reinvestment Act. Traditional analytics in healthcare have been largely retrospective analyses based on claims data that seek to reconstruct an accurate picture of past events. Technological advances have addressed some barriers, but the limitations of data structure and the ubiquity and volume of clinical information — defined as data from electronic medical records (EMRs) and other sources — have introduced new challenges in an era of big data. In a recent joint survey administered in partnership by the College of Healthcare Information Management Executives (CHIME) and eHealth Initiative (eHI), the major institutional obstacles to adopting analytics include: lack of standardized data across systems, lack of system infrastructure to support analytics, and the cost of analytic software, among others.² These obstacles have depressed adoption of data-driven decision-making among health systems despite organizations’ desire to increase their analytic capabilities.

The eHI survey also polled respondents on the use of analytics in the course of daily business. Only 54.1 percent of organizations use data for operating room analytics, 44.3 percent for supply-chain analytics, and 14.8 percent for accountable care operations. Conversely, the majority of organizations continue to use data for financial management (91.8 percent), hospital reporting (86.9 percent), and operational efficiency (86.9 percent).² Clinical outcomes measurement is trailing behind operational and financial analysis, which isn’t surprising given that clinical data sources are only now becoming available in a discrete and analytically accessible digital format. Hospital adoption of EMRs has more than doubled over the past the three years alone, thanks in part to the momentum provided by HITECH

² eHealth Initiative (eHI) and Healthcare Information Management Executives (CHIME) Survey, 2012

incentives. In 2009, only 13.4 percent of hospitals had adopted an EMR system.³ Today, the most recent numbers from the Office of the National Coordinator for Health Information Technology (ONC) suggest that 48 percent of eligible hospitals have made the shift to capturing their treatment data electronically.⁴ As EMR adoption and use rates continue to grow among provider groups and hospitals, the use of analytics will become more attainable.

Clinical data sets are emerging as analytic data sources, and healthcare analytics professionals can anticipate leveraging the real-time nature of this data to fire predictive algorithms as a part of a health system's business intelligence. According to eHI, the majority of provider organizations (58.3 percent) direct resources toward retrospective analytics. In comparison, only 16.7 percent do so for real-time decision support, 13.9 percent for optimization/efficiency, and 2.8 percent for predictive analytics.² While the majority of organizations are not utilizing prospective analytics today, the potential game-changing value can be inferred from the effect of prospective analytics on other industries. For example, Richard Fairbanks of Capital One transformed the credit card industry by leveraging consumer data to anticipate what prospective customers would look for in the ideal credit card product and rapidly developed offers that targeted their unique interests and needs.⁵ His innovation to target credit cards for individuals led to a revolution in credit card service delivery. If health systems can accurately predict future behavior and outcome patterns, the implications for clinical, operational, and financial decision-making will be immense.

Integrated clinical and administrative data streams will help provider organizations better understand what has happened in their daily operations and also determine point-of-care treatment decisions in anticipation of a patient's potential trends. These kinds of decisions require data, statistical knowledge to accurately interpret what the data is indicating, and cognizance of the inherent limitations.

The Institute for Healthcare Improvement's Triple Aim⁶ of healthcare improvement has been touted in literature and federal legislation as the national goal in reforming our health system. The Triple Aim framework consists of three dimensions for optimizing health system performance:

- Improving the patient experience of care (including quality and satisfaction)
- Improving the health of populations
- Reducing the per capita cost of healthcare

Various articles plumb the depths of the value derived from digital data capture and analytics to achieve the goal of simultaneously improving individuals' health and the patient experience while lowering the cost of delivering appropriate care. William Stead at Vanderbilt University⁷ explores this linkage in detail, and other analyses have documented outcomes from improving care through use of digital health records.⁸ The healthcare industry has slowly shifted from asking why we

³ www.healthit.gov/media/pdf/ONC_Data_Brief_AHA_2011.pdf

⁴ <http://www.healthcare-informatics.com/news-item/cms-onc-tout-milestone-ehr-adoption-rates-incentive-payouts>

⁵ <http://www.banktech.com/top-innovators/2011/Richard-Fairbank>

⁶ <http://www.ihl.org/offerings/Initiatives/TripleAim/Pages/default.aspx>

⁷ *Information Knowledge Systems Management 8 (2009) 119–143 119, DOI 10.3233/IKS-2009-0140, IOS Press*

should achieve the Triple Aim, to asking very detailed questions about how we can achieve it. With new sources of data available for integration and analysis, the possibilities to use information to help with this tectonic shift have now become reality.

Ten Observations From the Trenches

Truven Health AnalyticsSM has been engaged in healthcare data analytics since the origin of using administrative data to support financial and operational decisions in the 1980s. Truven Health began by unlocking the power of administrative data sources from unstandardized and unproven billing data and has evolved to provide scalable, statistically advanced, retrospective analytics to large organizations today. In parallel, Truven Health spent years working with hospital systems' clinical feeds to drive analytics at the point of care with alerting and rounding tools that help deliver real-time clinical interventions. The following discussion outlines insights from the trenches of integrating clinical and administrative data sources to drive new value for our clients and partners. While not exhaustive, our observations and approach may provide insights to other organizations wrestling with the complexities of integrated analytics.

1. Anticipate and Plan for Trust Issues When Accessing Multi-Stakeholder Data

Organizations have been eager to seize opportunities to expand the visibility of patients across the care continuum. Many business plans call for access to data from other providers, commercial payers, and other “non-owned” data sources. While this is an exciting goal, trust barriers are time-consuming and difficult to surmount. The legal and psychological fabric of trust will be established slowly, and implementation efforts predicated on external data can compromise launch timeframes.

Best Practice: Implementation of new analytic data streams requires two simultaneous work streams. The first track addresses the technical issues involved in connecting HL7, CCD, or other new inbound streams of data, contending with the typical challenges of multiple vendors and source systems. An equally important second track is the privacy and legal work stream comprised of acceptable and limited data-use rights for data sources, participation agreements, business associate agreements, and the thorough communication and vetting of plans with privacy officers and other stakeholders. As these other efforts mature, anticipate reasonable delays in external data inclusion and the beginning stages of analytics integration across the silos in your organization.

2. Accommodate Legacy Systems Rather Than Waiting for “Future Perfect”

Developing countries outpace U.S. mobile phone infrastructure in terms of speed and capabilities, largely as a result of the existence of legacy infrastructure.⁹ The U. S. health system (despite significant recent investment) has a similarly large stock of existing infrastructure developed prior to the era of exchange and analytics. Building analytic capabilities predicated only on future EMR capabilities, Integrating the Healthcare Enterprise (IHE) connectivity, and codified data elements will delay ROI from initial analytic insights and jeopardize support for future initiatives.

⁸ Cebul, R. et al. *Electronic Health Records and Quality of Diabetes Care. New England Journal of Medicine* 2011, pp. 825-833

⁹ http://money.cnn.com/2012/02/21/technology/spectrum_crunch/index.htm

Best Practice: Work with what exists today. Leverage HL7 2.x data streams, unstructured clinical notes, and existing claims sources as raw material for initial analytics. Engage with technology solutions that can intake, link, and analyze the imperfect current state while still offering capabilities to incorporate newer standards for interchange. A no-data-left-behind strategy best accommodates current realities, introduces analytics at a lower initial price point, and begins providing informational advantages while the underlying infrastructure is gradually upgraded.

3. Understand the Context of Data Creation

Most data used for clinical, operational, and financial metrics was not originally created with analytics in mind. While most information systems and analytic professionals are aware of this fact, the interpretation problems are multiplied when cross-integrating traditional claims and clinical data sources. For example, patient-classification data embedded in inpatient clinical data sources reflects the workflow of bed management and must be interpreted carefully as a roll-up for analytic classification when combined with episode-grouped claims sources. Likewise, the motivations of those involved in source data entry (charge entry, billing accuracy, or simply paging through admissions screens to complete a quick patient registration) greatly impacts the quality of specific fields used for subsequent analysis.

Best Practice: Employ an analytically aware integration of data streams into analytic repositories. Ensure that data management and integration professionals adequately understand the end goal of analysis so that appropriate transformations occur during data loading and editing. When deploying real-time analytics against unstructured data, ensure that dynamic ad-hoc querying is conducted in partnership with professionals who understand the source systems and the operational processes that populated them. This context preserves credibility in the face of clinical and leadership scrutiny of the resulting insights.

4. Adjust to Uneven Data Availability via Adaptable Analytic Methods

Analysts have awaited access to clinical data streams with great anticipation, only to be disappointed in the lack of consistency regarding data population and coding. Uneven implementation of medication administration record (MAR) systems, lack of standardization of biometric data elements, and other data gaps are common. Workflow and EMR development will continue, but the resulting capture of data beyond Meaningful Use-mandated data packages and partial streams will continue to be reality. Data analysts could spend an inordinate amount of time working with input professionals, various system vendors, and IT staff in a perpetual effort to move the data toward completeness and quality.

Best Practice: The healthcare industry can't afford to wait for complete and rational data sets. Deploy techniques and solutions that accommodate gaps in data. Impute trends statistically based upon best-available data, understanding the full data set denominators and full trend lines of data will rarely be complete across the continuum of care. Fundamentally, leverage analytic systems that have give, balancing analytic precision where possible with the realities of underlying data limitations. Ensure that analytic insights delivered preserve the understanding of these tradeoffs at the point of consumption.

5. Utilize Comparisons and Benchmarking

Thus far we have discussed access to owned and trusted stakeholder data sets. While this provides a certain level of analysis, the inevitable managerial question is, “What are my competitors doing?” Performance measurement and benchmarking activities are critical not only to an individual organization, but also the field at large as the healthcare system seeks to constantly improve. An empowered healthcare consumer movement has spurred activity to publicly report performance at an industry level and, at a federal level, to force-rank and evaluate hospitals. An organization that can rank itself against self-created benchmark targets at a local, regional, and national level will call attention to its strengths and have a better understanding of how to compete in the changing delivery market.

Best Practice: Partner with organizations that can provide statistically viable benchmarking capabilities. Ensure that these benchmarking data sets can also integrate with your analytics to provide a valid outside perspective at the local, regional, and national levels. Pair dynamic and ad-hoc analysis with benchmarking, especially when seeking behavior change and influence from physician and clinical leaders.

6. Account for Asynchronous Data Streams

Data is populated from clinical and administrative processes at different rates. Clinical data arrives in rapid-fire increments via workflow-initiated trigger events, backload from clinical data repositories, and health information exchange (HIE)-mediated events. Administrative data is often delivered in batches at the conclusion of billing cycles. This results in data availability ranging from “ticker-tape” style streams to monthly or quarterly uploads. The nearly constant flow of refreshed clinical data streams is sometimes at odds with the statistically significant sample sizes needed for the application of advanced analytic flags.

Best Practice: Utilize tools and systems that permit asynchronous data collection across multiple periodicities. Acquire capabilities that can apply analytic methods rapidly in a source-appropriate manner. Architect solutions that allow insights to be produced with partial data, refresh quickly over time, draw on statistically valid conclusions, and integrate the clinical and administrative data sources quickly and repeatedly.

7. Anticipate Lack of Data Standardization

Despite the work of numerous standards bodies and the progress of Meaningful Use, EMR data still has varying levels of terminology consistency. Continuity of Care Document (CCD) standards are defined, but unevenly populated and implemented in the practice of individual vendors. The breadth of previous visit patient history, for example, may only be selectively included in care transition data flows. Process-of-care measures beyond those mandated by Physician Quality Research Institute (PQRI) and other regulatory requirements are poorly structured or not documented at all. Where terminology coding has been applied to clinical information, it is often done inconsistently or not updated frequently enough. When building a longitudinal understanding of a person or a population across multiple data sources,

discrepancies quickly materialize. For example, a person's ZIP code registers as one location from an ADT (admission, discharge, transfer) feed and another location from a claims record. Which is correct for geographic roll-up purposes? And how does one determine this? Furthermore, some of the strongest insights reside in unstructured physician notes and transcribed report text strings that are often not included in analysis.

Best Practice: Incorporate the use of terminology standardization in the analytic process for measures that require strong retrospective statistical validation. Ensure that integration tools preserve and display source-system attribution so that data conflicts can be readily resolved (or at least understood). Deploy techniques such as natural language processing and semantic logic to mine unstructured text for value. Minimize integration and data management costs by not forcing data standardization on elements that are primarily used for directional or contextual analysis.

8. Be Practical In the Face of “Big Data”

Healthcare analytics organizations routinely work with large administrative claims data sets through state Medicaid agencies, large health systems, and CMS. The volume of data from clinical sources (representing an equivalent number of care episodes), however, dwarfs these databases. Truven Health recently analyzed a statewide all-payer billing database containing approximately 1 million records, representing four years of discharge history. In that same state, we performed an analysis on four months of EMR data for three medium-sized hospital systems that generated a data set of 2.8 million records. Population-level analysis can be hampered by performance and processing speeds if not anticipated and managed appropriately. The aggregation of millions of records to understand comparative outcomes of female hernia patients admitted on Tuesdays versus Thursdays could become an intense use of computing power for a potentially questionable inquiry. Organizations such as Google have solved this problem by engaging intense amounts of inexpensive processing power with multiple failover systems.¹⁰ Other options include the traditional strategy of creating summary roll-ups of certain measures stored separately in order to provide shortcuts to answering anticipated queries.

Best Practice: A practical approach to computational and data-mining techniques begins with thinking carefully about what one wants to ultimately measure. Terminology mapping and standardization should be selectively applied to data sets that will be intensively used. Eschew upfront efforts to aggressively codify all inbound sources of data in order to spread out integration costs and speed up implementation time. Utilize flexible data-mining techniques to sift through large and relatively unstructured pools of clinical data that will be used less intensively. Apply real-time and predictive analytics against data stored at the point of receipt, and apply large-scale retrospective analysis against standard measures and those data subsets that are pre-calculated to increase efficiency.

⁹ http://en.wikipedia.org/wiki/Big_data and <https://developers.google.com/appengine/>

9. Implement Analytics Incrementally

By its very nature, analytics answer one question only to uncover new questions for further inquiry. No vendor or analyst, no matter how prescient, can anticipate all the needs through pre-structured systems. Ad-hoc analysis through query development and reporting is an important capability, as is the ability to reframe data during the course of analytic discussions. Emerging market demand exists for both the roll-up of data to aggregate populations across the continuum of care, as well as the granular detail of drilling down into an individual patient record. As healthcare organizations move toward accountable and coordinated care paradigms, they will need administrative and clinical data that are integrated and linked by patient, provider, and episodes of care.

Best Practice: Use agile practices to develop your analytics roadmap. Begin with the most critical handful of questions to be answered and work backwards to the data sources necessary to support those analytics. Most organizations can only focus on a few operational levels at any point in time — determine the most likely candidates and begin integrating these data sets. Leave the no-data-left-behind concept, but build out analytics on top of those data sources as business needs dictate, and don't attempt to read the minds of your constituents.

10. Analytic Chain of Trust

An organization's analytic systems ultimately collapse without a complete analytic chain of trust. Confidence in data sources, derivation of the analysis, and those presenting the data are paramount. For example, analytics are increasingly being used to monitor behavior and performance and to determine (in part) compensation at a physician and executive level. The determination of targets and goals that impact shared savings and other risk-sharing contracts are sensitive, and it is crucial that they are accurate. As scientists by nature, physicians in particular will scrutinize data sets and underlying methodologies for accuracy and validity. Buy-in from stakeholders depends on an analytically aware implementation, a mastery of data sources, and methodologies and tools that account for imperfect data. Failure to cover these bases will impair organizational credibility and ultimately limit the use of data for decision-making. While our industry has HIPAA-focused security and privacy offers protecting data, equivalent scrutiny should be applied to the logic chain supporting analytic insights.

Best Practice: Ensure a complete understanding of the analytic chain of trust from data acquisition to ad-hoc querying, to terminology management, and data manipulation. Choose trusted and proven partners to appropriately curate and manage data assets, and provide benchmarking input and statistically sound methodologies adapted to new data sets and structures. Appreciate the limitations of each link in the chain. Once trust is broken, it is not easily recovered.

Moving Forward

Healthcare organizations will gradually build a comprehensive analytic understanding of persons and populations built on a foundation of integrated administrative and clinical data. Pragmatic data analytics, however, begin with an appreciation of the limitations of the current state. Organizations developing their integrated data analytics strategy should appreciate the uneven terrain of data sets and terminology, the trust fabric of inter-organizational data sharing, and the need for external benchmarking and analytic methods adapted to underlying data variability. The real-time, high-volume, and unstructured nature of emerging clinical data sets requires a flexible approach to data collection and analysis.

While integrated analytic solutions have not yet become ubiquitous in health systems' business strategies, they are gaining traction to play a critical role in helping providers and organizations improve the quality of care delivered. The recent eHI/CHIME survey is sobering in noting that although 93 percent of organizations agreed that data and analytics were very important to the future of their organization, only 28.2 percent believed they have what they need to meet their current analytic needs today.² However, with appropriate analytic partners and adoption of key best practices noted in this paper, healthcare organizations can reinforce the analytic chain of trust that is crucial to fulfilling their primary mission to patients and stakeholders.

About the Authors



Grant Hoffman

Vice President of Clinical Data Integration

Grant Hoffman is the Vice President of Clinical Data Integration for Truven Health Analytics, leading product management and service delivery for Health Information Exchange and clinical analytics solutions. Prior to assuming this role, Grant led service delivery functions for the hospital benchmarking solutions group. Grant joined Truven Health via the acquisition of MercuryMD, where he was the Vice President of Client Services. He has 18 years of healthcare information technology and services experience, including work with the initial stages of several technology startups. Grant also served as a healthcare information technology consulting manager with Ernst & Young LLP (now Accenture) and several roles with the McKesson Corporation. He received his bachelor's degree from the University of North Carolina at Chapel Hill and his MBA from the Kenan-Flagler Business School at the University of North Carolina.



Rebecca Molesworth

Manager of Data Integration Markets and Strategy

Rebecca Molesworth is the Manager of Data Integration Markets and Strategy for Truven Health Analytics, responsible for direct market intelligence functions for the data integration portfolio, including solution definition, packaging, and data integration strategy support. Prior to this role, Rebecca was the Program Manager for several emerging business initiatives at Truven Health, including Health Information Exchange, Mobile Health Information Exchange, and Payment Reform Analytics. Rebecca joined Truven Health from WebMD Health Services, where she was the Director of Sales for the employer market. Rebecca has 10 years of prior healthcare information technology experience, working with several condition management and behavioral wellness companies including WebMD, Miavita LLC, and Matria (now Alere). She received her MHS in Health Management and Policy from the School of Public Health at the University of Michigan.



FOR MORE INFORMATION

Send us an email at

info@truvenhealth.com

or visit truvenhealth.com



ABOUT TRUVEN HEALTH ANALYTICS

Truven Health Analytics delivers unbiased information, analytic tools, benchmarks, and services to the healthcare industry. Hospitals, government agencies, employers, health plans, clinicians, pharmaceutical, and medical device companies have relied on us for more than 30 years. We combine our deep clinical, financial, and healthcare management expertise with innovative technology platforms and information assets to make healthcare better by collaborating with our customers to uncover and realize opportunities for improving quality, efficiency, and outcomes. With more than 2,000 employees globally, we have major offices in Ann Arbor, Mich.; Chicago; and Denver. Advantage Suite, Micromedex, ActionOI, MarketScan, and 100 Top Hospitals are registered trademarks or trademarks of Truven Health Analytics.

truvenhealth.com | 1.800.366.7526

©2012 Truven Health Analytics Inc. All rights reserved. All other product names used herein are trademarks of their respective owners. HOSP 12021 1112