

Healthcare Edition

MELISSA DATA Magazine

Fraud, bad debt, duplicates, compliance...

Find the solutions to these challenges.

Rx

for the
Bad Data
Blues



- +** 4 steps to a better data warehouse
- +** A unique approach to turn duplicates into one Golden Record
- +** How CalOptima reduced costs with improved data quality

Cover Story



12

RX for the bad data blues

A 4-step plan to improve data warehousing and business intelligence efforts through better data

by Abby Telleria

Features

03

Why Melissa Data?

04

Application Story

There's Gold in Them Thar Records

Create a Golden Record from Duplicate Patient Records Using a Data Quality Score

by Joseph Vertido



09

Contact Data Quality Consulting

10

Case Study

CalOptima Migrates Clean Data to a Data Warehouse

Solutions Spotlight

14

Data Quality Tools

Personator, Global Address Verification, Record Matching & Deduplication, Geocoding, Change of Address

15

Database Administrator Tools

Contact Zone, Data Quality Components for SQL Server

Free Demo



Try out all our capabilities using our Listware for Microsoft Excel plugin. **Free test.**

www.MelissaData.com/healthcaremag

Melissa Data.

Your Partner for Better Healthcare Data.



Since 1985, data quality's been our obsession – the driving force behind the creation of many of our scalable data cleansing and enrichment solutions. Our data quality tools verify, standardize, consolidate, enrich, and update U.S., Canadian, and international contact data, like address, name, phone, and email information. What sets Melissa Data apart is the multitude of reference data sets at our disposal – including property, consumer, business, demographic, firmographic, and international data sources that propel your business forward.

At Melissa Data, clean data is in our DNA. We understand that your patient data is a precious commodity. That's why we strive to help healthcare providers and organizations achieve the highest quality data to improve data management, HIPAA compliance, patient outcomes, and billings and claims processing.

More than 10,000 companies rely on Melissa Data to gain a single, accurate, and trusted view of critical information assets.

For all your data quality needs, [call 1-800-MELISSA](tel:1800MELISSA) or [visit www.MelissaData.com](http://www.MelissaData.com) for an office near you. Our offices are located in Rancho Santa Margarita, CA; Rockwall, TX; Braintree, MA; Berlin, Germany; London, UK; and Bangalore, India.



Rx for the Bad Data Blues

A 4-step plan to improve data warehousing and business intelligence efforts through better data

By Abby Telleria, Senior Writer

Healthcare organizations struggle with data quality – from serious issues involving fraud, bad debt, billing inefficiencies – to life-threatening issues involving diagnosis and prescribing care and real-time analytics.

As healthcare organizations try to handle increasing amounts of data – coming from disparate entry points with inconsistent data standards, to siloed and/or legacy data – all while working to transition from paper to electronic health records (EHR) – many healthcare organizations are choosing a data warehousing solution that integrates data quality into all of their applications to maintain the accuracy and value of business and critical-care operational information.

See the next page for the four recommended steps to eliminate the root causes of data problems before they negatively affect your data warehousing and BI efforts.





Fraud costs the healthcare industry nearly **\$70 billion** annually.
-National Health Care Anti-Fraud Association (NHCAA)



U.S. hospitals experienced **\$41 billion** in uncompensated care.
-American Hospital Association



Bad data costs organizations between **15 - 20 percent** of their operating budgets.
-U.S. Insurance Data Management Association



Accurate patient data could prevent **90 percent** of claim denials.
-U.S. Insurance Data Management Association

Step 1 >> Profile Your Data

The adage “garbage in, garbage out” applies more today than ever before as data-centric systems play an increasing role in supporting healthcare decision making. Data quality issues involving deduplication, incompleteness, and inconsistencies all play a part in undermining the effectiveness of operations.

Data profiling is often considered the starting point in the data quality process. It's the critical phase in assessing and analyzing the quality of data values within and across data sets.

Data profiling helps provide a thorough understanding of data quality issues. Without data profiling, it is almost inevitable that development costs will spiral and projects will overrun, or fail outright.

Step 2 >> Verify, Correct, Update and Standardize Your Data

Today's healthcare industry is complex, data-driven, and tightly regulated. It's essential to identify your organization's data issues and determine the most effective approach to cleaning and updating your data.

One major data issue affecting healthcare organizations - outdated data. Contact data is always in flux; patients are constantly moving, changing jobs, retiring, getting new phone numbers, etc. As a result, healthcare organizations face problems associated with undeliverable mail, which impacts the efficiency of billings and collections, and creates costs associated with returned mail and wasted postage.



Quality information is essential to all aspects of today's healthcare system, so improving the quality of data, information, and knowledge is paramount as we transition from paper to EHRs.”

- American Health Information Management Association (AHIMA)



Every year around 43 million Americans move. Utilize **SmartMover® Change-of-Address processing service** from Melissa Data to regularly update patient address information and keep in contact with patients that have moved, while reducing undeliverable mail and costs associated with wasted printing and postage.

Melissa Data's **Personator** cross-references millions of records from a multi-sourced database to confirm all of the contact information, including address, name, phone, and email address provided relate to a specific person. Name-to-address verification helps certify the identity of an individual to minimize risks associated with bad debt and fraud.

Invalid or inaccurate contact data can also enter systems when data entry personnel makes a typo or misspelling during the admission process. And, patients are also responsible for erroneous or even false information entering systems when they fill out registration forms, which leave healthcare organizations susceptible to insurance fraud and bad debt. Sometimes, patients might not provide their most current contact information – either intentionally or unwittingly. Other patients might not give their first name, preferring to list only their nickname or middle name. The most common types of fraud involve false statements or deliberate omissions that are critical to determine healthcare eligibility or for billing purposes.

The impact of fraud is huge – costing the healthcare industry nearly \$70 billion annually, states the National Health Care Anti-Fraud Association (NHCAA). “This loss directly harms patients, taxpayers, and the government through higher healthcare costs, insurance premiums, and taxes.”

The key to solving these issues is to verify patient data at the point of care and keep it updated throughout the billing process to help ensure timely payments and reimbursements.

Cross referencing all contact data associated with a patient, including verifying the patient actually lives at the address provided, can help prevent fraudulent activity and mitigate risk.

Developing rules and deploying a data quality regimen for patient contact information will result in a cleaner, more accurate database that helps meet regulations, including HIPAA compliance, achieve a single view of the patient, and improve communications between providers and patients.



STEP 3>> Fill in the Gaps of Missing Information

Healthcare organizations can enhance the value of their database by adding additional pieces of information to improve communications, get a more complete view of patients, and for business intelligence. Adding missing email addresses, phone numbers, and completing an address with the correct suite or apartment number empowers omnichannel communications that increase patient engagement, lower costs, improve quality, and helps retain patients and members. Adding demographic and/or geographic data can help identify quality disparities and other patterns of care.



STEP 4>> Merge, Eliminate and Consolidate Duplicate Records

Duplicate medical records are an alarming and costly threat to the healthcare industry as they can negatively impact patient safety, hospital liability, reimbursement, and administrative efficiencies. Duplicate records can also undermine the integrity of a data warehouse.

American Medical Informatics Association (AMIA) cited a report by Fox and Sheridan that stated an average organization's duplicate rate is between 5 to 10 percent for a single hospital.



Melissa Data's **Personator** also adds missing data elements such as names, phone numbers, and email addresses to help ensure accuracy and improve communication efforts.

Melissa Data's **Geocoding** solution assigns precise latitude and longitude coordinates to domestic and international addresses to empower business intelligence based on location. Available in Personator and Data Quality Suite.



92 %

of duplicate errors occur during in-patient registration.

-Johns Hopkins Hospital



Identifying duplicates is a critical component for a high level of data precision. Melissa Data's **MatchUp®** weeds out duplicates and merges multiple records into a single, accurate view of the patient – the Golden Record – through a process known as survivorship.

The report also estimates that a duplicate pair of records creates \$50 in hidden operational costs, so a hospital that generates five duplicates a day could end up spending as much as \$78,000 per year as a result of duplicate records. If the hospital is open seven days a week, costs rise to over \$91,000 per year according to the AMIA.

Duplicates are Dangerous

Not only are duplicates costly, they can be harmful to patients. A provider could mistake one patient for another patient with a similar name, especially if there are duplicate records of the patient in the system. For example, a “Beth Smith” might be recorded as “Smith, Elizabeth” in another database, but both names are the same person. Or, the provider might associate “Beth Smith” and a “Beth Smithe” as the same patient, but they are actually two different people.

In another instance, a provider might miss critical information on a patient – like allergic reactions to certain drugs, or if a patient was recently admitted to the hospital for another illness, etc. – because that patient had multiple records in the database. In essence, one record might indicate this vital information, but the duplicate record might not – leading to a potential misdiagnosis or wrong treatment plan.

Detecting and consolidating duplicate records into a single, accurate view of the patient helps prevent unnecessary costs, streamline databases, improve patient care and operational efficiencies, and gives providers a more holistic view of patient data.

Keeping Data Healthy for the Long Term

By implementing these four steps into a data quality regimen – (1) Profile, (2) Verify and Update, (3) Enrich, and (4) Consolidate – healthcare providers can integrate patient information from a range of varied sources to build and maintain an effective data warehouse that provides a more accurate view of patients, improves communications, and ultimately, helps improve care and patient outcomes.

Where to Go for Answers



When healthcare is data driven, patient data is interconnected and available when and where it's needed to save lives. As data becomes knowledge, medicine becomes more precise. However, the data needed for analysis and operations is often of poor quality.

Since 1985, Melissa Data has offered flexible, scalable, powerful data quality tools to maintain the trustworthiness of data, correct errors, and eliminate duplicate records. To help you implement and optimize the value of these tools, we offer Contact Data Quality Consulting. Our consulting services help you identify data quality issues and develop a plan to ensure timely and accurate updates to contact information so your data warehousing efforts are protected and bad data doesn't affect the decision-making process.

Here's how our data quality consulting can help your organization:

- **Improve** data collection with point-of-entry verification
- **Meet** regulatory requirements, including HIPAA compliance
- **Decrease** bad debt by maintaining accurate records for billing/reimbursement
- **Consolidate** records for a single view of the patient

Talk with us today about your unique operational and technical needs and together we'll work to ensure healthier outcomes and a healthier bottom line.

Call 1-800-MELISSA



About CalOptima

CalOptima is a county organized health system providing publicly - funded health coverage programs for low-income families, seniors, and people with disabilities in Orange County, California.



CalOptima Migrates Clean Data to a Healthcare Data Warehouse



• Challenge

CalOptima serves 422,000 members with a network of more than 5,800 primary care doctors and specialists, as well as 24 hospitals. As the second largest health insurer in Orange County, California, CalOptima provides coverage to one in seven residents, and one in three of the county's children.

Deploying a Data Warehouse

In order to better serve its members, improve communications, and grow its business, CalOptima embarked on a tremendous challenge – migrate its membership data into a data warehouse.

A key issue CalOptima faced was the poor quality of its member data, specifically inaccurate mailing and address



Our main objective was to add value to the data produced for business users. Having conformed and integrated data is valuable, but the added value of recognizing valid addresses is significant.”

**-Osvaldo Cruz,
CalOptima’s data warehouse architect**



information. CalOptima wanted to reduce undeliverable-as-addressed mail and the costs associated with returned mail.

“Our main objective was to add value to the data produced for business users. Having conformed and integrated data is valuable, but the added value of recognizing valid addresses is significant. One of the cost-driven indicators is returned mail and missing communication with members,” Osvaldo Cruz, CalOptima’s data warehouse architect said.

“Since CalOptima is a member-centric organization, adding quality around our ability to communicate with members is crucial to support our mission,” Cruz added.

• Solution

Data Quality Components for SQL Server Integration Services (SSIS) are data transformation tools available from Melissa Data that provide a full complement of data verification, standardization, and consolidation functionality.

• Results

With so much data to collect and manage, CalOptima recognized it needed to implement a solution that

would clean and standardize all of its multi-sourced, multi-dimensional membership data – a critical step during migration to a data warehouse. That is why the insurer turned to Melissa Data for its Data Quality Components for SQL Server.

“It is very easy to integrate with new and existing ETL components,” Cruz said. “The fact is, a self-contained data flow component makes its implementation an easy task, taking full advantage of the SQL Server Integration Services engine.”

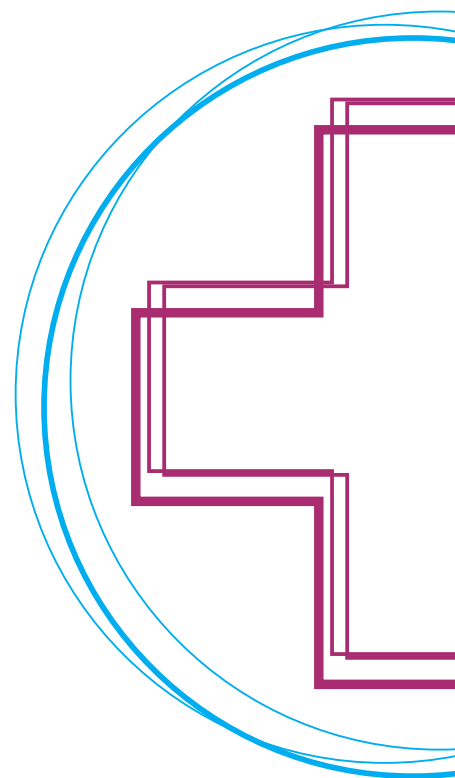
Cruz noted that it was the component’s ease-of-use that lured them to the product. “We appreciated the extensive functionality provided in the out-of-the-box version. It is an excellent, self-documented product, and has a very intuitive user interface for programmers and architects. It is a simple drag-and-drop to include SSIS components in the ETL flow.”

CalOptima was also able to take advantage of the solution’s geospatial functionalities. Now the organization can determine the distance between their members and their primary care providers.

“The additional geospatial attributes can leverage the existing BI platform in place, in order to produce accurate

maps and other geographical representations ... so we now have better knowledge of our members’ locations,” he said.

Cruz expects that the solution’s biggest impact will be on reducing returned mail costs.



THERE'S GOLD IN THEM THAR RECORDS:

Create a Golden Record from Duplicate Patient Records Using a Data Quality Score

By Joseph Vertido, Data Quality Analyst

The importance of survivorship, known as the “Golden Record” in data terms, is quite often overlooked in the quest for clean and effective patient data. Yet this final step in the record matching and consolidation process is more important than ever in business today, it ultimately allows for the creation of a single, accurate, and complete version of a patient record.

Even if you have invested in a state-of-the-art matching tool, and through careful analysis have constructed a matching policy that will catch all the duplicates in your database – how do you determine the most accurate data to use in establishing the Golden Record? Applying intelligent rules based on reference data – rather than just using the most recent record – is a new approach, and one that is increasing the value of Golden Record data.

➤ Golden Record Basics ➤

After running the matching process, you may be presented with the duplicated records bundled nicely into duplicate groups and ready for consolidation. Obvious matches such as John Smith at 123 Main St and John Smythe at 123 Mein Street, are both identified as having the same information. What comes next? What do you do with the duplicates once they are detected?

In the matching methodology, choosing the unique or winning Golden Record is the next logical step. The process of selecting surviving records means selecting the best possible candidate as its representation.

However, “best” in the perspective of survivorship, can really mean a lot of things. It can be affected by the structure of data, the source of the data, how the data is populated, what kind of data is stored, and sometimes by the nature of business

rules. Consequently, techniques can be applied in order to accommodate certain types of variations when performing survivorship.

➤ Traditional Survivorship Techniques ➤

Which record do you keep as your survivor and which ones do you discard? There are three commonly used techniques in determining the surviving record. In the most recent methodology, date-stamped records can be ordered from most recent to least recent. The most recent record can be considered eligible as the survivor.



The most frequent approach matches records containing the same information as an indication of their correctness. Repeating records indicate that the information is persistent and therefore reliable. Finally, the most complete method considers field completeness as its primary factor of correctness. Records with more values populated for each available field are considered the most viable candidates for survivorship.

Although these techniques are commonly applied in survivorship schemas, correctness may not be as reliable in many circumstances. Because these techniques apply to almost any type of data, the basis on which a surviving record is created conforms only to “generic” rules. In contrast, by leveraging reference data, database administrators (DBAs) can generate better and more effective schemas for survivorship.

➤ Evolving to Reference Data ➤

Applying reference data in survivorship changes how rules come into play. Using the most recent, most frequent, or most complete logic really has more of an aesthetic basis for selection.

Ideally, the selection of the surviving record should be based on an actual understanding of data.

And this is where reference data has impact. Most importantly, it focuses solely on being able to consolidate the best quality data. By incorporating reference data, DBAs gain an understanding of the actual contents of data and create better decisions for survivorship. Changing the perspective as to how the quality of data is defined, in turn, breaks the norm of typical survivorship schemas.

Survivorship Decisions

Let's take a look at some instances involving how reference data and data quality affect decisions for survivorship.

I. Address Quality

Address quality is essential, and separating good data from bad data should take precedence in survivorship decisions. In the case of addresses, giving priority to good addresses makes for a better decision in the survivorship schema as opposed to selecting the most frequent.

Name	Address	City	State	Zip	Status
John Doe	22382 Avenida Empresa	RSM	CA	92688	Valid
John Doe	22382 Empresa	RSM	CA	92688	Invalid
John Doe	22382 Empresa	RSM	CA	92688	Invalid

MOST FREQUENT

II. Record Quality

It could also be argued that good data may exist in a single group of matching records. In cases like these, you can assess the overall quality of data by taking into consideration other pieces of information that affect the weight of overall data quality. Take for example the following data:

Name	Address	City	State	Zip	Phone	Email
	22382 Avenida Empresa	RSM	CA	92688	Valid	
John Doe	22382 Empresa	RSM	CA	92688	8008006245	
John Doe	22382 Empresa	RSM	CA	92688	111111111	johndoe@ya

MOST COMPLETE

VALID PHONE NUMBER

In this case, the ideal approach is to evaluate multiple elements for each record in the group. Since the second record contains a valid phone number, it can be given more weight or more importance than the third record, despite the third being more complete.

These images illustrate that the methodologies and logic used for record survivorship become dependent primarily on data

quality, whether you're working with contact data, product data, or any other form of data. The focus on data quality transcends, and even overrules, other determining factors, such as which record was most complete or latest.

For example:

Name	Address	Zip	Phone	Date
John Doe	22382 Avenida Empresa	92688	8008006245	01/01/2005
John Doe	22382 Avenida Empresa	92688	111111111	01/01/2013

It can be argued that the second record is the most recent one, and should therefore be the survivor. But upon careful consideration of the quality of the data, you can see that the second record contains an invalid phone number.

This type of intelligent approach enables businesses with a more human perspective and allows them to correctly conclude that the first record has the better data and should therefore be established as the Golden Record.

Applying New Perspective

However you choose to define data quality, it is imperative that you keep only the best pieces of data if you are to have the most complete information. The most powerful future for data quality lies in the new and unique ability to discern contact data quality information and select the surviving record based on level of quality of the information provided.

This new technique for Golden Record selection offers a more effective and logical approach when it comes to record survivorship. Ultimately, this creates an automatable system that can make better decisions for data cleansing – creating a single, accurate, high-value version of the truth that actually makes business sense.

The Golden Record selection is currently available within the MatchUp SQL Server Integration Services (SSIS) Component. For more info and a free trial, go to:

www.melissadata.com/data-quality-ssis/matchup-component.htm



Joseph Vertido is a data quality analyst and MVP channel manager at Melissa Data. He can be reached at joseph@melissadata.com

Solutions Spotlight// At a glance

Test Our Tools for Free



Our tools are available as multiplatform APIs, in the cloud, or as enterprise solutions. Request a free trial of any tool or easily test all our capabilities in Microsoft Excel through our Listware plug-in. **Get a free trial now!**

www.MelissaData.com/healthcaremag



Data Quality Tools



Personator®

Melissa Data's flagship data quality cloud service, Personator, helps healthcare providers make clear connections within and across data sets, to resolve different representations of a record, and link all touch points of contact data together to achieve a single, accurate, and complete view of their patient.

Personator Verify

Name-to-address-verification:

- Compares data with multisourced data sets to determine accuracy
- Verifies name, address, phone number, and email address all correlate to same individual
- Geocode address to rooftop level (lat/long)
- Move Update (10+ years)

Input:

Joseph Vertido
22382 Avenida Empresa
Rancho Santa Margarita, Ca
92688-2112
joseph12@gmail.com

Output:

Address
22382 Avenida Empresa
Rancho Santa Margarita, Ca
92688-2112

Name

Joseph Vertido

Email

joseph12@gmail.com



Personator Append

Completing your records:

- Add missing email, phone numbers, names, and company names to your records
- Demographic Data

Input:

Joseph Vertido
22382 avenida empresa
92688

Output:

Joseph Vertido
22382 Avenida Empresa
Rancho Santa Margarita, CA
92688-2112
949-589-5200
joseph12@gmail.com



Data Quality Suite (World & North America Editions)

The Data Quality Suite verifies domestic and international address, email, name and phone information, and provides precise latitude and longitude coordinates. The suite helps maintain accurate contact data to improve mail deliverability, lower costs, and minimize bad debt.



MatchUp®

MatchUp is a powerful advanced matching tool to help identify the most difficult-to-detect duplicate records – giving your organization the most accurate, unified view of your patient information. By correctly consolidating your data using MatchUp, you will reduce costs and dramatically improve staff efficiencies and patient care.



SmartMover®

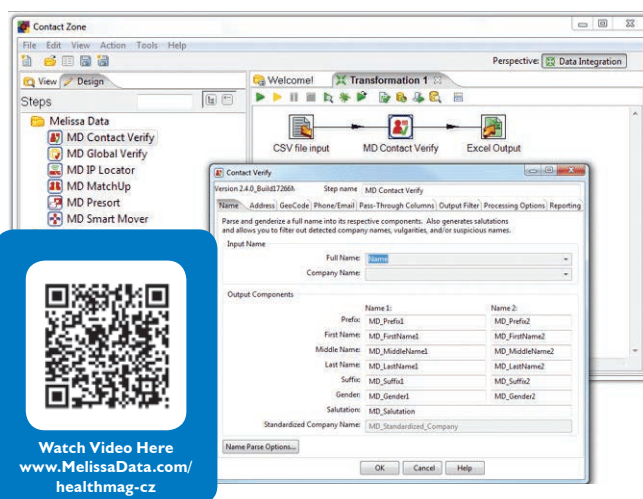
Use Melissa Data's NCOALink® service to get the updated addresses of your existing patients and/or customers and reduce wasted materials and postage associated with undeliverable-as-addressed mail. Our service compares your mailing list records to the 48 month USPS® NCOALink database. Melissa Data also offers Canadian National Change of Address processing.



Database Administrator Tools

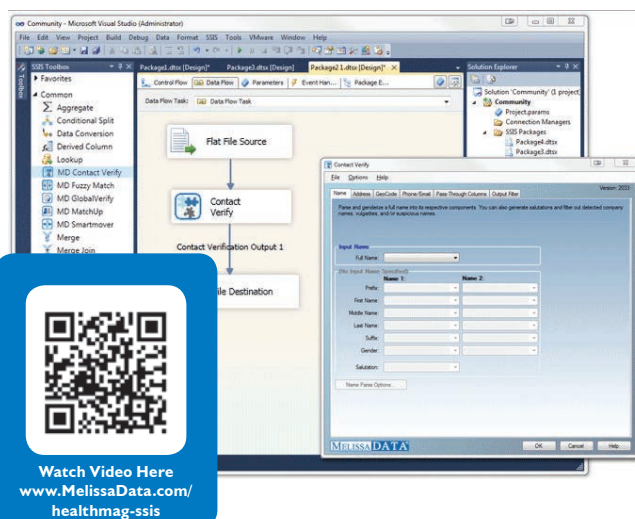
Contact Zone®

Melissa Data's Contact Zone makes it easy for any healthcare organization to clean, validate, integrate, and enrich patient information – without the need for extensive programming skills commonly required to implement a data quality solution. The software can run several of Melissa Data's products including Personator, MatchUp, Geocoding, and Presorting.



SQL Server Integration Services (SSIS)

For healthcare providers who prefer to use the Microsoft® platform to cleanse their database, Melissa Data offers its Data Quality Components for SQL Server Integration Services. This suite of easy-to-use data transformation components for Microsoft SSIS verifies, corrects, standardizes, updates, and consolidates patient data.



MELISSA DATA CORP.
22382 Avenida Empresa
Rancho Santa Margarita
California, 92688-2112

To correct or remove your name from this list, fax your request to 949-589-5211

Better data means better healthcare.

Poor quality data can adversely affect every aspect of the healthcare payment and delivery system. Since 1985, Melissa Data has provided thousands of organizations with the data quality tools needed to clean, update, consolidate, and enhance patient, provider, and member data to improve care and lower costs.

Helping you create a blueprint for improved patient care.



MELISSA DATA®

Get a free trial at
www.MelissaData.com/healthcare
1-800-MELISSA