

# Collect SMART

---

Project Manager Quick Start Guide

THE NORTH CAROLINA  
Institute *for* Public Health

Developed by the North Carolina Institute for Public Health

*v November 6, 2015*

Development Team:

Principal Investigator: Rachel Wilfert, MD, MPH, CPH

Project Manager: Matt Simon, MA, GISP

Research Associate: John Wallace, PhD, MSPH

Research Associate: Kasey Decosimo, MPH

Curriculum Developer: Meredith Davis, MPH

Curriculum Developer: Kristen Ricchetti-Masterson, MSPH

Acknowledgements:

David Zane, Tracy Haywood, Dr. Jennifer Horney and her PHEB 689, Methods in Field Epidemiology class, epidemiology graduate students from Gillings School of Global Public Health and community members, staff, and volunteers from Durham County Department of Public Health, Pender County Health Department, Chatham County Public Health Department, Forsyth County Department of Public Health, Gaston County Department of Health and Human Services, and Granville-Vance District Health Department for providing feedback during software piloting.

For more information about this project, contact:

North Carolina Institute for Public Health

Gillings School of Global Public Health

The University of North Carolina at Chapel Hill

Campus Box 8165

Chapel Hill, NC 27599-8165

919-843-5561

[collectSMART@unc.edu](mailto:collectSMART@unc.edu)

<http://sph.unc.edu/nciph/ta/SMART>

# How to Get Started with Collect Smart

To get started with Collect SMART you will need to install both desktop and mobile software and establish an account at the Collect SMART website. Then, in order to begin using the software for your first project, you will need a set of spatial data files and a digital questionnaire form. First, we'll cover how to get the software up and running. You will need the following 1) desktop version of Epi Info™ with Collect SMART encryption keys, 2) an account set up at [collectsmartdata.com](http://collectsmartdata.com) and 3) the Collect SMART android app. Details on each step are described below followed by the spatial data requirements and the digital form.

## Software set-up

### Step 1: Desktop Epi Info™ 7

Collect SMART is designed to be used with the Centers for Disease Control and Prevention's Epi Info™ 7, a suite of software tools for your PC that allows for easy form design, data entry and analysis. Users should use the *Create Forms* module in Epi Info™ to create a data collection form and then export this form as a template (.xml). This .xml file will be uploaded to Collect SMART when you start a new project.

Collect SMART utilizes the Collect Data module of Epi Info™ Companion for Android and generates two sets of files when users capture survey response; an .xml and an encrypted and password protected sync file (.epi7). In order to import these sync files you must download and install the Collect SMART version of Epi Info™ 7 [HERE](#) that will read the encryption keys created by Collect SMART. The current build of Epi Info™ is 7.1.5.2. Form design can occur in the CDC's version of Epi Info™ which can be downloaded [HERE](#). However, Sync files created by Collect SMART can only be imported using the Collect SMART version of Epi Info™.

To install Epi Info Desktop follow these steps:

1. **Download** the zip file from dropbox by clicking this link: [Collect SMART version of Epi Info™](#).
2. Using compression software like WinZip or WinRAR, extract the Epi Info™ files. Choose a location close to your root such as C:\software\EpiInfo.
3. Locate the **Launch Epi Info7.exe** file and double click to run. If you can't run the .exe, open a command prompt to run the file.
  - a. To find the command prompt click the start button, and search for 'command prompt.'
  - b. With the command prompt open change your directory to the location of the Launch Epi Info 7.exe file. To do this type 'cd C:\software\EpiInfo' using the example directory above.
  - c. Type 'Launch Epi Info 7.exe' into the prompt and hit return.
4. To save time when re-opening, pin the .exe to your start menu or task bar

## Step 2: Account set-up at CollectSMARTdata.com

To use Collect SMART you must establish a user account at the website <http://collectsmartdata.org/>

1. Choose the **SIGN UP** option in the upper right hand corner of the home page.
2. Provide your **full name**, **email address** and a **user name** and **password**. This user name is the **coordinator** name used to log into projects on a mobile device.
3. Click **Submit**

Please allow up to 5 days for your account to be approved. If you do not receive a confirmation email within 5 business days send an email to [CollectSMART@unc.edu](mailto:CollectSMART@unc.edu). Once you receive the confirmation email you will be asked to click a link to validate your account and then you can use your login and password to login to the site.

## Step 3: Installing the Collect SMART app for Android

In order to install an app that does not come directly from the Android Play store, you must copy the .apk from your computer to the device and tell your device to trust apps from 'unknown sources.'

1. You must first enable developer options. For most Android devices this is under **Settings -> About phone/device or just About**.
2. Next locate the **Build number** and tap this **7 times**. After two taps, a small pop up notification should appear saying 'you are X steps away from being a developer' with a number that counts down with every additional tap.
3. After the 7<sup>th</sup> tap the developer options will be unlocked and available. Find and open these options. Usually they are on the main Settings menu.
4. Next **Enable USB debugging**. This will allow you to copy files directly from your PC to your device.
5. Copy **CollectSmart-v0.9.62.apk** from your PC to the **Download** folder on your device.
6. Navigate to the Download folder and tap on the .apk file to install. After installation hit done and restart your device before opening Collect SMART for the first time.
7. When you login to Collect SMART on your device for the first time you will need to supply the **username**, **password** and **coordinator**. The coordinator name is the user name you used to set up an account at CollectSMARTdata.com. Before you can create user accounts you must create a project, which will be described later in the guide.

If you are installing a new version of the app follow these instructions:

1. Uninstall the previous Android app using Settings —> Apps. If using Android 4.4 or 5.0.2 then make sure to use the options menu (top right corner of screen "...") to select "uninstall for all users". Confirm that app was uninstalled by scrolling through the "Settings —> Apps —> All" list to make sure that Epilinfo is not part of the alphabetically ordered list or at the end of the list.
2. Use the ESFileExplorer app to navigate to the "download" folder. Then either delete or rename the "Epilinfo" folder. To select the folder do a long press on the folder name.

3. Connect the table to a PC and make sure that USB Debugging is enabled on the tablet. Copy the new apk to the Download folder and install.
4. Restart device and open app.

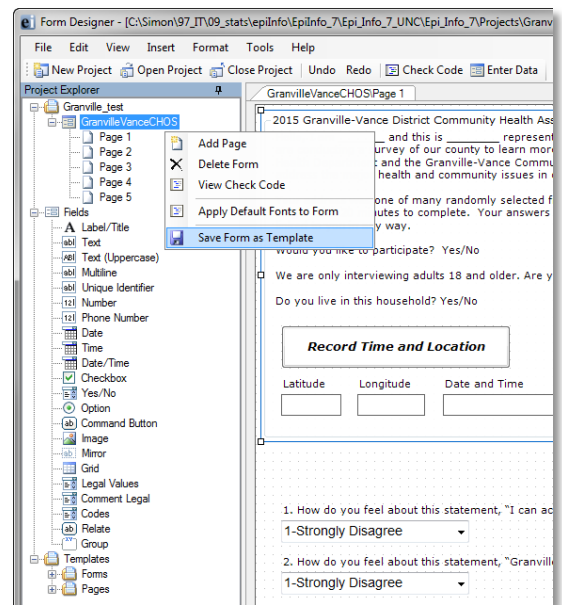
## Required spatial and data files

Collect SMART was designed for cluster sampling, but can accommodate other sampling methods such as simple random sampling or convenience sampling. For all projects the first step is to design a digital data entry form in Epi Info™ (Step 1: xml creation). For cluster sampling you need a file showing the geographic regions that have been selected for the survey. If you are conducting a two-stage cluster sample you should also have a set of survey points (Step 2: spatial data files).

### Step 1: .xml creation

Forms should be coded using the Form Designer in the Create Forms module of Epi Info™ 7. Once you have added a couple fields and coded some questions, it is a good idea to test out your form for errors and to make sure any check code you may have written is functioning as intended. Also, **always** be sure to test out your finalized form before live data collection.

To test your form, the form you need to create an .xml file. To do this, right click on the form name and choose **'Save Form as Template'** (see image to the right). This will create an .xml in the \Epi\_Info\_7\Templates\Forms directory. Copy this .xml over to your project directory. To make things easier, create an 'upload' folder that has your .xml and the .kmls which will be covered in the next section.



When coding forms in Epi Info keep in mind there are some limitations to the current version of software. The most commonly used field types are **text, legal values, comment legal** and **checkbox boxes**. You can use other field types such as date/time and number fields, but be sure they are displaying properly on the mobile device AND storing the data correctly. Currently, the option buttons are not storing data properly in the .epi7 file, though the data is captured correctly in the xml. We are actively working to resolve this issue.

Another thing to keep in mind is that the real-time analysis widget is still beta and relies on parsing logic to populate itself. This parsing is affected by labels in ways that we are exploring further. Use labels sparingly. Refer to the example xml to see what a successfully coded form looks like. (survey\_Granville\_CHOS\_2015\_FINAL.xml). Currently, only legal and comment legal values are being parsed correctly by the analysis widget. We are actively working on correcting this.

**Avoid using special characters such as '&', '<' and '>.'** These special characters create a .epi7 file that can't be imported into Epi Info desktop. This is a limitation of the current release and will be fixed with the next major release.

## Step 2: spatial data files

You have two options when using Collect SMART. You can supply your own spatial data or you can have Collect SMART create the necessary files.

### Supplying your own data

Use this option if you have used a GIS site selection tool to create your sample or have a set of randomly selected locations identified ahead of time. First you must get your files formatted correctly. Collect SMART only accepts **.kml** files with a specific set of properly ordered and named fields. In order for the Collect SMART to properly parse your .kml files you must include the fields mentioned below or the app will not work properly.

#### **Cluster.kml**

"<Placemark id=" or  
"<PLACEMARK\_ID="

"<name>" or "<NAME>"

"REF\_CENT\_X" or "Ref\_Cent\_X" or  
"ref\_cent\_x"  
"REF\_CENT\_Y" or "Ref\_Cent\_Y" or  
"ref\_cent\_y"

"<COORDINATES>" or "<Coordinates>" or "<coordinates>"

#### **Points.kml**

"<PLACEMARK ID=" or "<Placemark id=" or "<placemark id=" or  
"<NAME>" or "<Name>" or "<name>"

"POINT\_X" or "Point\_X" or "point\_x"  
"POINT\_Y" or "Point\_Y" or "point\_y"  
"STREETADD" or "Streetadd" or "streetadd"  
"CITY" or "City" or "city"

Note that for some fields there is flexibility in lowercase and uppercase. The ordering of the REF\_CENT\_X and REF\_CENT\_Y fields in Cluster.kml can be switched, but all other fields must maintain the order specified above. Note that the field names surrounded by < > are automatically created when you create a .kmz using the Layer to KMZ tool in ArcMap. However, you must be sure to **label** your points and clusters in ArcMap. This will place the label into the <Name> field. <COORDINATES> and <Placemark id> is automatically generated during the layer to kmz conversion. The field names that are highlighted are the ones that you must have in your ArcMap layer before converting.

Collect SMART HOME CONTACT US LOGOUT

## Welcome to Collect SMART

The power of Collect SMART sampling is in its ability divide the population of interest into randomly-selected geographically distinct clusters using U.S. Census blocks or block groups. The Collect SMART sampling tool will draw a random sample with probability proportional to population based on the number of housing units in your sampling frame. The sampling frame can be defined as a single county, multiple counties or by zip code or multiple zip codes.

Within the clusters, a random sample of households should be selected to be interviewed. Refer to the User Guide for Project Managers on acceptable second stage sample selection methods. This method has been validated for rapid assessments and the estimation of population-level health needs (producing valid and precise estimates of +/-10%).

Before you create your sample you will need to create your questionnaire using the desktop version of Epi Info 7. You can [download that here](#)

If you only want to use Collect SMART to manage your random sample survey (or "convenience survey") go here to select a geographic region and upload your questionnaire.

Or if you only want to use Collect SMART to manage your survey and you have already pulled your sample go here and upload your spatial data and your questionnaire.

**CLUSTER**

**CONVENIENCE**

**UPLOAD FILES**

Collect SMART creates spatial data

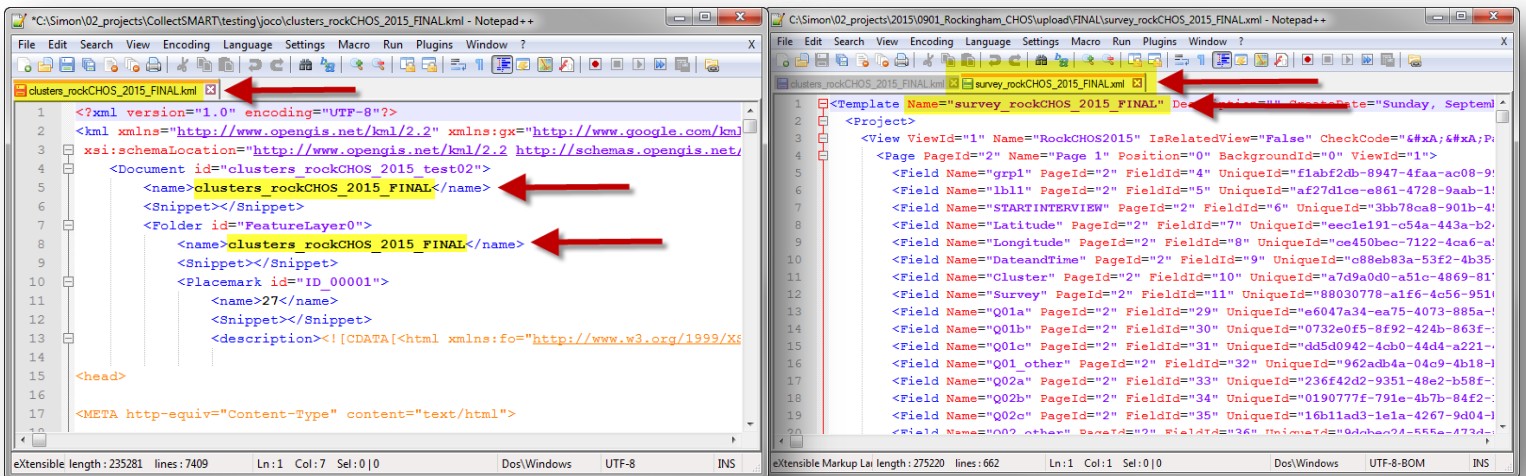
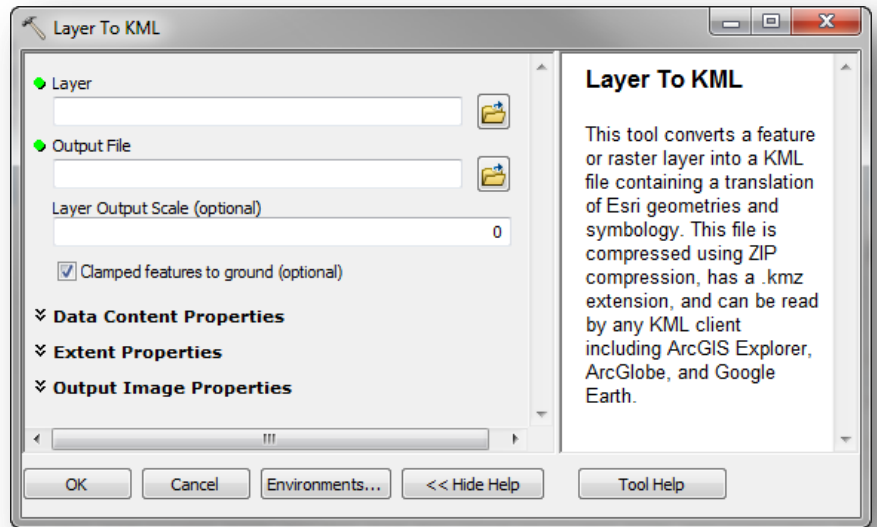
Supply your own data

UNC THE NORTH CAROLINA Institute for Public Health SKY HIGHWAYS Mobile Software epi info

Also, in the Points.kml file if the <NAME> value = "HOTEL or "BASE" then we use this entry as the Staging Area.

Once you have created a .kmz you need to change the extension name to .zip and use compression software to unzip the file. This will create a .kml file that you can rename. Collect SMART requires that the 'Survey Name' be in each of the files that you upload and the cluster, point and questionnaire file be named in a specific manner. For example, if you were conducting a survey in Gaston County in 2015 you might enter "Gaston\_2015" for the survey name. Then the cluster name would need to be named "clusters\_Gaston\_2015", your points need to be names "points\_Gaston\_2015" and the questionnaire should be named "survey\_Gaston\_2015."

The final step is to make sure the internal name of each file matches the filename described above. To do this you will need a text editing software such as Notepad++, which is a free source code editor that supports different programming languages and includes syntax highlighting. Open your spatial files (.kml) and your form (.xml) and make sure the name field matches the name of the file exactly. Be careful not to edit any other text with the editor open and be sure to save the changes.



Once your spatial files are properly named with the correct fields and you have a form coded (.xml), you are ready to upload these files to Collect SMART.

From the Collect SMART home page choose **Create** and then choose **Upload Files**. First type in your **Survey Name** and specify the state and date. This survey name must be unique and can't be reused. If you enter a name that has been used in the past, you will be notified and prompted to try another name. Then choose your cluster, point and questionnaire file and click **Upload Files**.

**Collect SMART** HOME CONTACT US LOGOUT

### Upload New Survey

Please complete all the following fields then hit "Upload Files" button:

- Enter New Survey Name:(up to 64 characters)
- Select Survey State :
- Enter Survey Date
- Cluster File:  No file chosen  
Cluster file needs to be a .kml formatted file with the following fields:  
Placemark id (or PLACEMARK\_ID)  
name (or NAME)  
REF\_CENT\_X (or "Ref\_Cent\_X" or "ref\_cent\_x")  
REF\_CENT\_Y (or "Ref\_Cent\_Y" or "ref\_cent\_y")  
COORDINATES (or "Coordinates" or "coordinates")  
You can download [AN EXAMPLE CLUSTER.KML FILE HERE](#)
- Points File:  No file chosen  
The point file also needs to be a .kml format file with the following fields:  
PLACEMARK ID (or "Placemark id=" or "placemark id=")  
NAME (or "Name" or "name")  
POINT\_X (or "Point\_X" or "point\_x")  
POINT\_Y (or "Point\_Y" or "point\_y")  
STREETADD (or "Streetadd" or "streetadd")  
CITY (or "City" or "city")  
Also in the Points.kml file if the "NAME" value is "HOTEL" or "BASE" then this entry is assumed to be the Staging Area.  
You can download [AN EXAMPLE POINTS.KML FILE HERE](#)
- Questionnaire Filename:  No file chosen

UNC UNIVERSITY OF NORTH CAROLINA  
THE NORTH CAROLINA Institute for Public Health  
sky highways Mobile Software  
epi info



## Using Collect SMART to supply spatial data

Collect SMART requires spatial data for cluster sample and convenience samples. If you are conducting a convenience sample then you must first supply a **project name**, **date**, **state**, your sampling frame (which can be multiple counties or zip codes) and a **staging area**. Make sure to tab out of the staging area box and to properly geocode the address. When successful, you will see the address and coordinates appear below the box.

**Create New Project**

Use this page to generate a convenience sampling frame specified by one of more counties or zipcodes

Please Enter New Project Name:(up to 64 characters)

Please Enter Survey Date

**Project generation steps :**

1. SELECT REGION (STATE/COUNTY OR ZIPCODE)	2. GENERATE SAMPLE FRAME	3. UPLOAD PROJECT FILES							
<p>State : <input type="text" value="North Carolina"/></p> <p>Select County/Countries :</p> <ul style="list-style-type: none"><li>Wake County</li><li>Nash County</li><li>New Hanover County</li><li>Northampton County</li><li>Onslow County</li><li><input type="text" value="Orange County"/></li><li>Pamlico County</li></ul> <p>Select Zipcode(s) :</p> <ul style="list-style-type: none"><li>27006</li><li>27007</li><li>27009</li><li>27011</li><li>27012</li></ul> <p>Please Enter Staging Area Address [street, town]: <input type="text" value="400 Roberson Street, Carrboro, NC"/></p> <p>400 Roberson St, Carrboro, NC 27510, USA [35.9094827, -79.0697955]</p>	<p>Selected:</p> <table border="1"><thead><tr><th>Region</th><th>Population</th><th>HUs Tracts</th></tr></thead><tbody><tr><td>Orange County</td><td>133801</td><td>55597</td><td>28</td></tr></tbody></table> <p>Enter the Maximum Number of Surveys/Points you wish to collect : <input type="text" value="100"/></p> <p><input type="button" value="GENERATE SURVEY"/></p>	Region	Population	HUs Tracts	Orange County	133801	55597	28	
Region	Population	HUs Tracts							
Orange County	133801	55597	28						

Next, specify the maximum number of surveys you plan on conducting. This number is used as a denominator when tracking survey progress and it populates the database with empty records. You will not be able to collect more surveys than this maximum, so be sure to pick a sufficiently large number such as 1000. Then click **Generate Survey**. This will create the sampling frame (referred to as clusters) and the empty points. Step 3, the final step, select **Upload Questionnaire** and navigate to your .xml, and then **Click here to complete Survey**.

The other option is to use the selection algorithm built into the platform top randomly selects blocks or block groups using the most current U.S. Census data. From the home page select **CREATE** and the page that says "Welcome to Collect SMART" choose **NEW PROJECT**. Before proceeding with the sample selection you should have your questionnaire designed in Epi Info™ and exported as an .xml, ready for upload to Collect SMART.

On the new project page (see figure below) you should complete the following: (1)Assign your project a name and (2) survey date, (3,4) define your sampling frame by choosing your state then choose either one county or multiple counties or zip codes, (5) choose the number of clusters you would like to select (6) choose the number of points or surveys you would like to collect in each selected cluster, (7) establish a threshold for the minimum number of housing units for a cluster to be included in the sampling frame, and (8) choose whether to use blocks or block groups.

Note that the Total number of surveys will be calculated and adjusts to your input. Finally choose (9) generate clusters to pull your sample. You will see a progress bar on the screen and then **Clusters complete** when the sample has been drawn. Once your sample has been generated, you can preview the clusters in Google Earth or some other GIS software that can read .kml files. If satisfied, upload your questionnaire file and then **Click here to upload Survey**. Once the survey project has been created you cannot edit the files.

**Collect SMART** [HOME](#) [CONTACT US](#) [LOGOUT](#)

## Create New Project

Use this page to generate a random sample of clusters in a sampling frame specified by one of more counties or zipcodes

Please Enter New Project Name:(up to 64 characters)  **1**

Please Enter Survey Date  **2**

**Please select sites for survey from the following options:**

SELECT REGION (STATE/COUNTY OR ZIPCODE)	GENERATE CLUSTERS								
State : <input type="text" value="North Carolina"/> <b>3</b> Select County/Countries : Dare County Davidson County Davie County Duplin County <input type="text" value="Durham County"/> <b>4</b> Edgecombe County Select Zipcode(s) : <input type="text" value="27006"/> <input type="text" value="27007"/> <input type="text" value="27009"/> <input type="text" value="27011"/> <input type="text" value="27012"/>	Selected: <table border="1"> <thead> <tr> <th>Region</th> <th>Population</th> <th>HUs</th> <th>Tracts</th> </tr> </thead> <tbody> <tr> <td>Durham County</td> <td>267587</td> <td>120217</td> <td>60</td> </tr> </tbody> </table> Enter the Number of Clusters you wish to generate (default 30): <input type="text" value="30"/> <b>5</b> Enter the Number of Surveys/Points you wish to generate within each Cluster (default 7): <input type="text" value="7"/> <b>6</b> <input checked="" type="checkbox"/> Require minimum of <input type="text" value="7"/> Housing Units per cluster <b>7</b> <input type="checkbox"/> Use Block Groups <b>8</b> Total number of Surveys = 210	Region	Population	HUs	Tracts	Durham County	267587	120217	60
Region	Population	HUs	Tracts						
Durham County	267587	120217	60						
To select multiple options: Hold down the control (ctrl) button and click (For Mac: Hold down the command button and click) <b>9</b>	<input type="button" value="CLICK HERE TO GENERATE CLUSTERS"/>								

### MANAGING PROJECTS AND USER ACCOUNTS.

The last step is to create a set of user accounts to access your newly created project. From the home screen of CollectSMARTdata.com, click the **MANAGE** button and then select **Manage User Accounts**. From this screen you click **Add User**, supply a *Username*, *Password* and the *Survey Name* from the previous step and then click **Submit** (Note that if you supply an incorrect project name, after you click submit the project name field will be blank). Use these login credentials when you open Collect SMART on your mobile device.

When a user has entered and saved a survey, a sync file will be automatically sent to the to the server. You can access these files by clicking on **Manage Projects** from the MANAGE screen. From

here you can delete you survey project (note this is permanent and cannot be undone), view your files and print cluster maps. When you choose **view survey files**, you will see all of the project files associated with the project including the sync files (.epi7) and xml files uploaded by users. At the end of your data collection, right click and save as the most recent sync file to import into Epi Info. For a quick look at the data you can click no the xml file in most browsers (we recommend Google Chrome).