

Figure A: Health Plan Data Collection Methods – Direct Primary Data Collection

METHODS	DESCRIPTION	ADVANTAGES	DISADVANTAGES
DIRECT PRIMARY			
<i>Enrollment</i>	Health plans can voluntarily collect race, ethnicity and preferred language information during the enrollment process.	<ul style="list-style-type: none"> • Data are self-reported and therefore fairly accurate. • Information collected at enrollment flows through to the rest of the data system. 	<ul style="list-style-type: none"> • Potential members may perceive that race, ethnicity and preferred language data might be used to deny coverage. • Health plan staff are not available to encourage members to respond and provide assurances. • If this is the sole method, used, it provides only one opportunity to collect information from members.
<i>Disease Management (DM) Programs</i>	Health plans can collect race, ethnicity and primary language information from their members as they enroll in targeted disease management or as case managers conduct interviews or outreach to enrollees. Method can also reach most vulnerable populations.	<ul style="list-style-type: none"> • Data are self-reported and therefore fairly accurate. • Provides multiple opportunities for data collection through frequent contact with member. 	<ul style="list-style-type: none"> • Data collected through DM programs may not necessarily be transmitted to other information systems within the health plan. • Data would be collected only for the subset of members who are in the DM program.
<i>Health Risk Assessments</i>	Health plans use health risk assessments (HRAs) to identify the future health care needs of their members. HRAs can collect data on race, ethnicity and primary language of a member.	<ul style="list-style-type: none"> • Data are self-reported and therefore fairly accurate. 	<ul style="list-style-type: none"> • Data collected through HRAs may not necessarily be transmitted to other information systems within the health plan. • Only captures data for those who complete the HRA form.

METHODS	DESCRIPTION	ADVANTAGES	DISADVANTAGES
<p><i>Encounter – Office Visit, Hospital Visit, etc.</i></p> <p><i>(May be considered a direct secondary source depending on the health plan.)</i></p>	<p>Medical groups, physicians’ offices, group practices and hospitals collect background information on patients during the admissions or intake process and these data may include race, ethnicity and preferred language.</p>	<ul style="list-style-type: none"> • Members have the opportunity to ask questions about why data are being collected and what data will be used for. • If staff are properly trained, this method can be quite effective in collecting data. • If a health plan is an Integrated Delivery System, shared systems and data infrastructure allow for easy data transfer from providers. 	<ul style="list-style-type: none"> • Data are typically not standardized or consistently collected by the different providers. • Providers or staff may be hesitant to ask these questions of patients, fearing litigation exposure. • Unless asking for self-identification, providers or staff noting/judging patient’s race or ethnicity may do so incorrectly. • Data is collected only for the subset of members who have an office visit or are hospitalized. • For network based plans, transmission of data may be difficult since there is no field for this type of information in claims forms.
<p><i>Health Plan Direct Contact</i></p>	<p>When a health plan sends an explanation of benefits to the members, it can include a survey or form asking for members’ races, ethnicities and preferred languages.</p> <p>Health plans cans also use interactive voice response (IVR) outreach calls to educate members about a topic and ask for a members’ race, ethnicity and preferred language.</p>	<ul style="list-style-type: none"> • Data are self-reported and therefore fairly accurate. • For minimal cost, this method can be paired with existing contact or outreach. • No additional data entry is necessary if using IVR. • Telephone-based outreach may be effective to address low literacy concerns. 	<ul style="list-style-type: none"> • Return rates for a survey or call completion of an IVR outreach call may be low. • Some types of outreach may not be an appropriate venue for collecting race/ethnicity and language data. Plans need to carefully consider the context for the outreach.

METHODS	DESCRIPTION	ADVANTAGES	DISADVANTAGES
<i>Member Web Portal</i>	Health plans can gather background information, such as race, ethnicity and preferred language, from members through their Web portals.	<ul style="list-style-type: none"> • Data are self-reported and therefore fairly accurate. • Allows for the collection of more granular data. • Eliminates the need for additional data entry. • Information would need to be collected only once. 	<ul style="list-style-type: none"> • Data is collected only for the subset of members who use the portal. • Plans need to be aware of the potential biases that are associated with Internet use. The data may over-represent certain subgroups.
<i>Member Satisfaction Surveys</i>	Health plans can use satisfaction surveys to collect data on race, ethnicity, and primary language. (e.g., Consumer Assessment of Health Providers and Systems (CAHPS) survey developed by AHRQ and NCQA).	<ul style="list-style-type: none"> • Data are self-reported and therefore fairly accurate. • A system is needed for transferring race, ethnicity and preferred language information from member satisfaction surveys to member files. 	<ul style="list-style-type: none"> • Data is collected only for the subset of members who respond to the survey. • Plans need to be aware of the potential biases that are associated with survey response. The data may over-represent certain subgroups.
<i>Member Contact (e.g., complaint or grievance, questions on benefits, or other administrative issues)</i>	At the end of the call, health plans can ask members to “update” their information. Information to be updated could include the member’s race/ethnicity or preferred language.	<ul style="list-style-type: none"> • Data are self-reported and therefore fairly accurate. 	<ul style="list-style-type: none"> • Not all types of member contact are an appropriate venue for collecting race/ethnicity and language information. Customers calling with grievances or complaints are less likely to cooperate. • Data are collected only for the subset of members who contact the health plan.

Figure B: Health Plan Data Collection Methods – Direct Secondary Data Collection

METHODS	DESCRIPTION	ADVANTAGES	DISADVANTAGES
DIRECT SECONDARY			
<i>Centers for Medicare & Medicaid Services (CMS)</i>	Plans that have a Medicare product can obtain these data from CMS. The Medicare program maintains beneficiary race and ethnicity data, derived from Social Security’s administrative records.	<ul style="list-style-type: none"> • Usually easy to obtain from CMS. 	<ul style="list-style-type: none"> • Accuracy of data may vary. • There are usually only four fields for data collection– white, black, other and unknown. • There are no separate fields for ethnicity. • Uses may be limited given lack of granularity.
<i>State (Medicaid)</i>	Information on an individual’s race, ethnicity and preferred language is collected during eligibility determination or enrollment in a health plan. States are required by CMS to identify the race, ethnicity and primary language of each Medicaid enrollee.	<ul style="list-style-type: none"> • Data can usually be easily obtained from the state. 	<ul style="list-style-type: none"> • Data sources, frequency of collection and accuracy vary significantly across states. • Need to establish a system for ensuring that monthly updates do not overwrite additional race/ethnicity and language information obtained by the plan.
<i>Employer</i>	Some employers collect race, ethnicity and language data for Equal Employment Opportunity purposes.	<ul style="list-style-type: none"> • The data often already exists through employer records. • Data are typically self-reported. • This method provides an opportunity for future collaboration between the health plan and employer. 	<ul style="list-style-type: none"> • Negotiating partnerships with employers can be complicated. • Employers with third-party administrators must negotiate data stream to health plan. • Employers may not collect data on dependents.

Figure C: Health Plan Data Collection Methods – Indirect Data Collection

METHODS	DESCRIPTION	ADVANTAGES	DISADVANTAGES
INDIRECT			
<p><i>Third-Generation Methods (Bayesian)</i></p>	<p>Bayesian methods use prior knowledge to evaluate the likelihood a member belongs to a particular group.</p>	<ul style="list-style-type: none"> • This method provides improved estimates from geocoding and surname analysis. • Can supplement direct data collection. 	<ul style="list-style-type: none"> • Data are not self-reported and therefore, may be less accurate.
<p><i>Combined Geocoding/ Surname Analysis</i></p>	<p>Geocoding provides estimates of the racial/ethnic composition of the surname area. Used in combination, these approaches can complement one another; geocoding can indicate race, and surname analysis can indicate ethnicity.</p>	<ul style="list-style-type: none"> • This method provides improved estimates from geocoding or surname analysis alone. • Can supplement direct data collection. 	<ul style="list-style-type: none"> • Data are not self-reported and therefore may be less accurate.
<p><i>Geocoding or Surname Analysis Alone</i></p>	<p>Geocoding is a method in which information about the social characteristics of the neighborhood or community a person lives in is used to infer information about that person (such as race, ethnicity or preferred language).</p> <p>Surname analysis uses a person’s last name to estimate the likelihood that he or she belongs to a particular racial or ethnic group.</p>	<ul style="list-style-type: none"> • This method can be can be a good place to start to supplement direct data collection. 	<ul style="list-style-type: none"> • Data is not self-reported and therefore, may be less accurate.