The Development of a New Tool for Data Quality Improvement and Use of an Immunization Registry in the 21st Century

Deborah Walker, PhD
Citywide Immunization Registry
New York City Department of Health
e-mail: dwalker@dohlan.cn.ci.nyc.ny.us

Vikki Papadouka, PhD
Citywide Immunization Registry
New York City Department of Health

Noam Arzt, PhD
HLN Consulting, LLC
e-mail: arzt@hln.com

Presented at:

National Immunization Conference

34th Meeting
July 5 - 8, 2000
Washington, DC

Background

New York Citywide Immunization Registry (CIR)

- Mandatory reporting started January 1, 1997 by Amendment to New York City Health Code
- All providers must report immunizations given to children under age 8
- 95% of institutions and 86% of private providers are currently reporting
- 100% of institutions and 92.5% of private providers have ever reporting
- Contains records for 1.8 million children with over 9.9 million immunizations
- Population based; Vital Records loaded every 2 weeks
- De-duplicating of records a high priority

Objectives

To describe a new query tool "Smart Search" which searches the CIR for possible matching records to an input file of children.

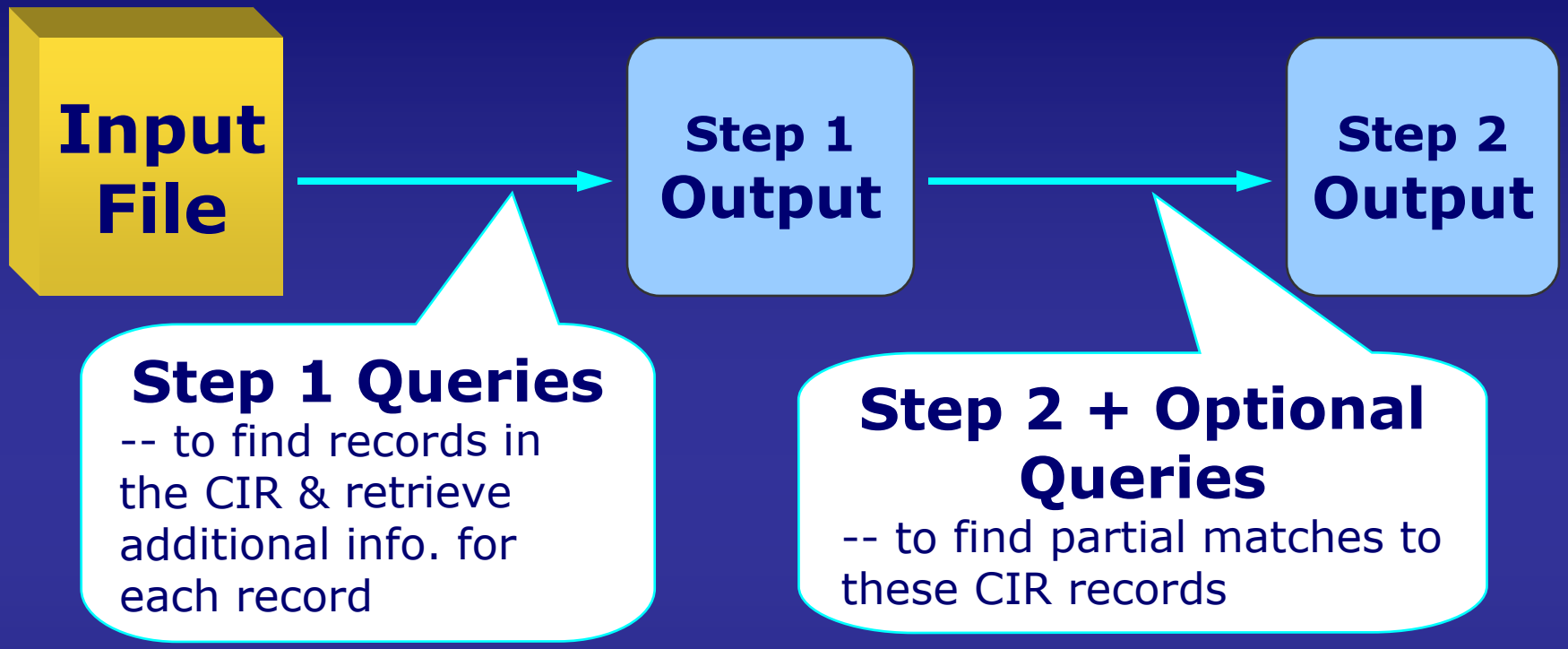
Smart Search

- A menu-driven query program to search CIR for possible matches
- User can select from over 1400 different combinations of search queries
- Used on Windows 95/98/NT desktops with log on access to CIR
- Developed in collaboration between NYC-DOH and HLN Consulting, LLC.
- Written in modules; additional functions and components easily added

Smart Search Functions

- Searches CIR for possible matches to an input file of records
- Creates file of possible matching pairs to input to **MEDD** (de-duplication program)
- Creates file to input to **Recall Generator** (reminder/recall program)
- Creates file to input to **DEI** (data exchange interface program) to exchange data with MCO's and NYC's Lead Quest Registry
- Generates patient lists and ASCII files of CIR records

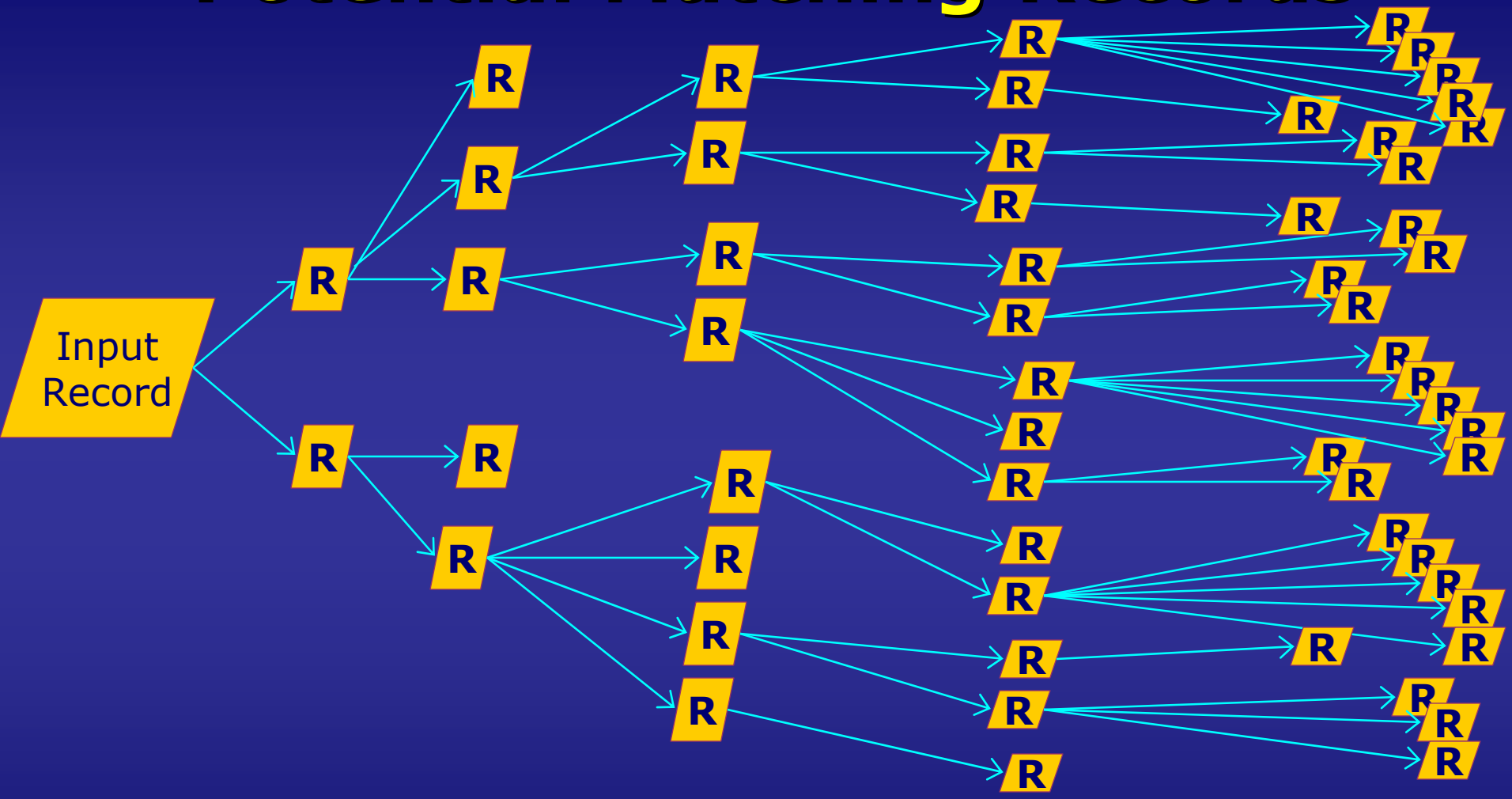
Smart Search Process



If *iterative searching*,

Process repeats Step 2 + Optional Queries up to 5 times.

Potential Matching Records



For every single record in Input File, multiple possible matches can be found depending upon search criteria.

Step 1 Output

Input file:

Papadouka, Vikki, 01/01/1999, F

Step 1 Output file:

```
248513512,"PAPADOUKAS","VASSILIKI",01/01/1999,F,"VP02135A","","123","","","","","ALEXANDRA",01/01/1930,1111111111,N,N
```

Step 2 Output file:

```
248702951,"PAPAD","VASSILIKI",01/01/1999,F,"","","","","","","","","",,,N,N
248702387,"PAPADOUKA","VASSILIKI",01/13/1999,F,"","","","","","","","","",,,
248702736,"PAPADOUKA","BABY GIRL",01/01/1999,F,"","","","","","","","","",,,
248702512,"PAPADOUKA","VIKKI",01/01/1998,F,"","","123","","","","","","",,,N
248702624,"MATTHEWS","VASSILIKI",01/01/1999,F,"VP02135A","","123","","","","","PAPADOUKA",01/01/1930,,N,N
248703063,"PAP","VASSILIKI",01/01/1999,F,"","","","","","","","","",,,N,N
248791375,"SPENCER","SHIRLEY",02/09/1999,F,"","","123","","","","","","",,,N
119145036,"WALKER","NATALIE",05/31/1993,F,"","","123","","","","","","",,7186360881,N,N
46370687,"HUIE","SHIRLEY",09/30/1996,F,"","","123","","","","","","",,,N,N
```

Step 1 Output

Input file:

Papadouka, Vikki, 01/01/1999, F

Step 1 Output file:

248513512
PAPADOUKAS
VASSILIKI
01/01/1999
F
VP02135A
123
ALEXANDRA
01/01/1930

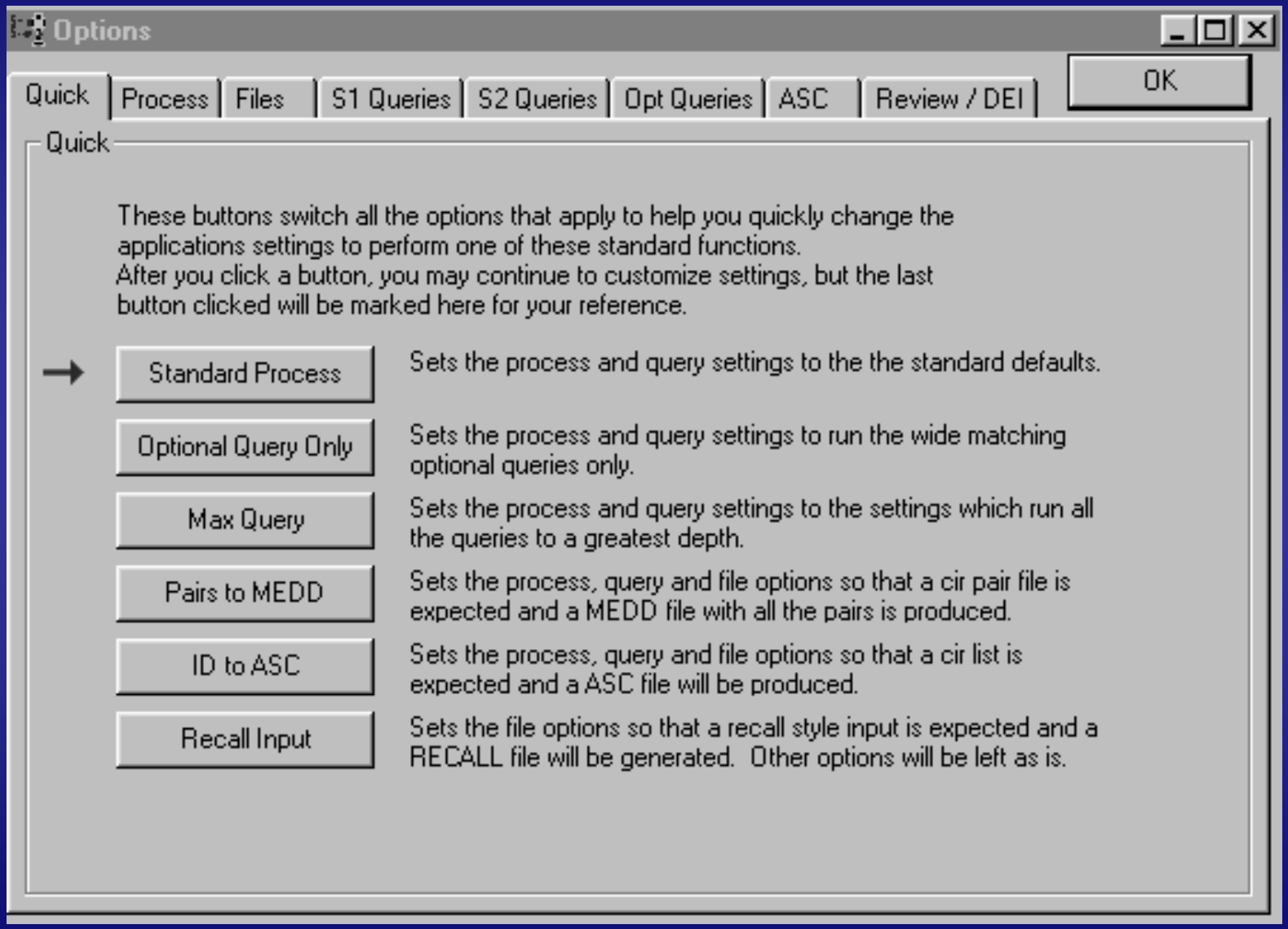
CIR Pairs file:

248513512 - 248702951
248513512 - 248702387
248513512 - 248702736
248513512 - 248702512
248513512 - 248702624
248513512 - 248703063
248702951 - 248791375
etc.

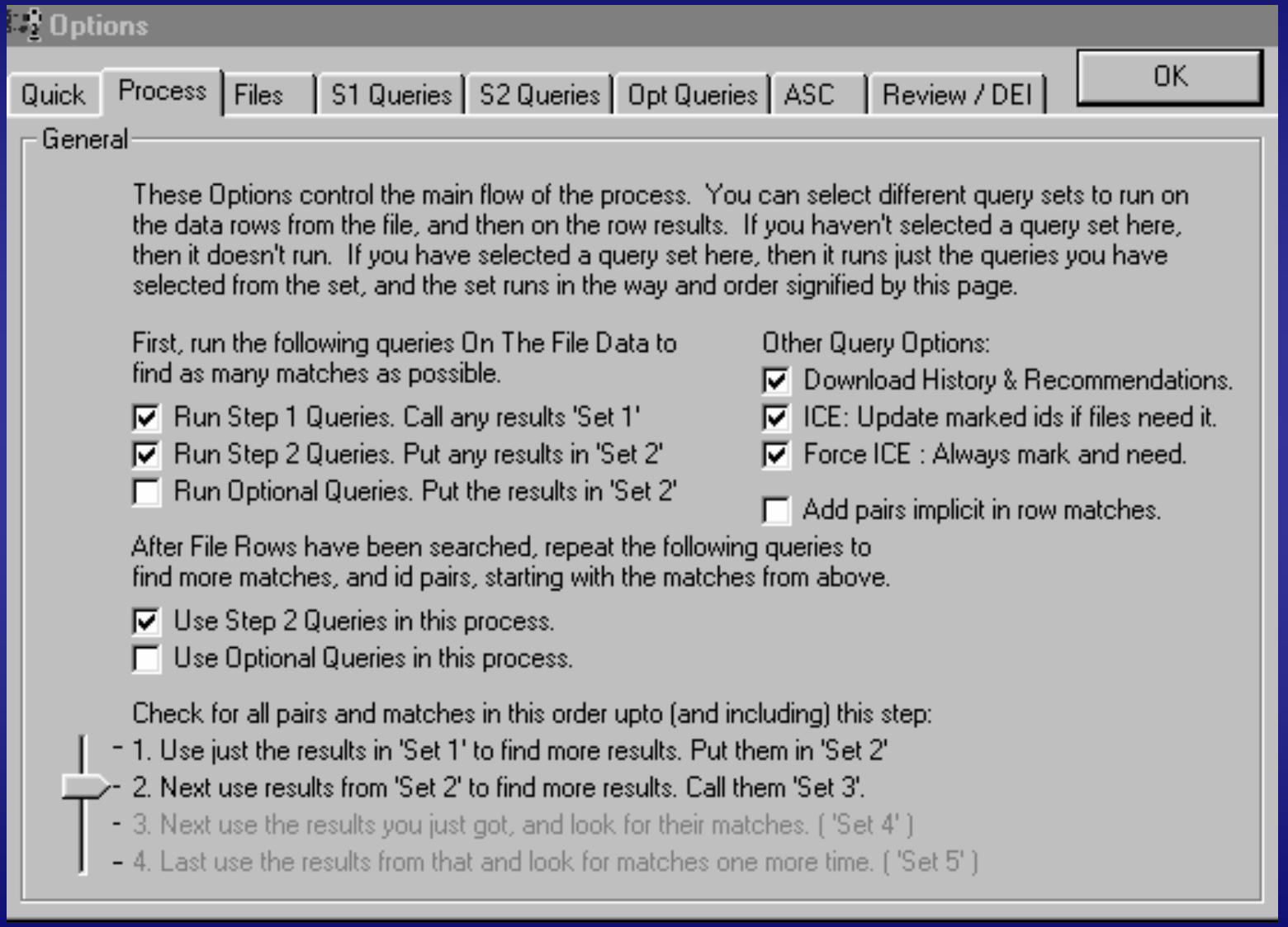
Step 2 Output file:

248702951 PAPAD VASSILIKI 01/01/1999 F	248702387 PAPADOUKA VASSILIKI 01/13/1999 F	248702736 PAPADOUKA BABY GIRL 01/01/1999 F	248702512 PAPADOUKA VIKKI 01/01/1998 F 123	248702624 MATTHEWS VASSILIKI 01/01/1999 F VP02135A 123 PAPADOUKA 01/01/1930	248703063 PAP VASSILIKI 01/01/1999 F	248791375 SPENCER SHIRLEY 02/09/1999 F 123	119145036 WALKER NATALIE 05/31/1993 F 123 7186360881	46370687 HUIE SHIRLEY 09/30/1996 F 123

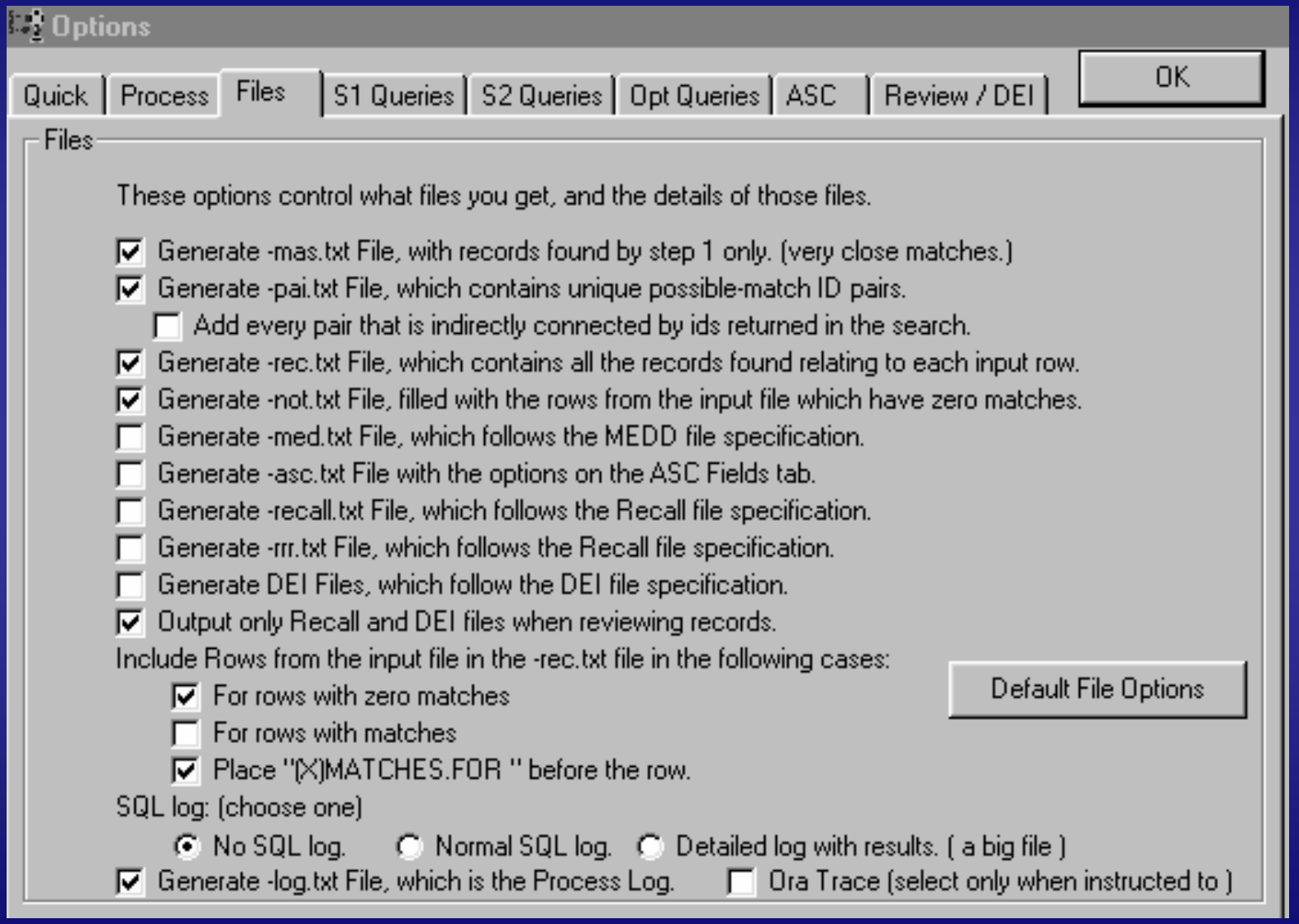
Smart Search Functions



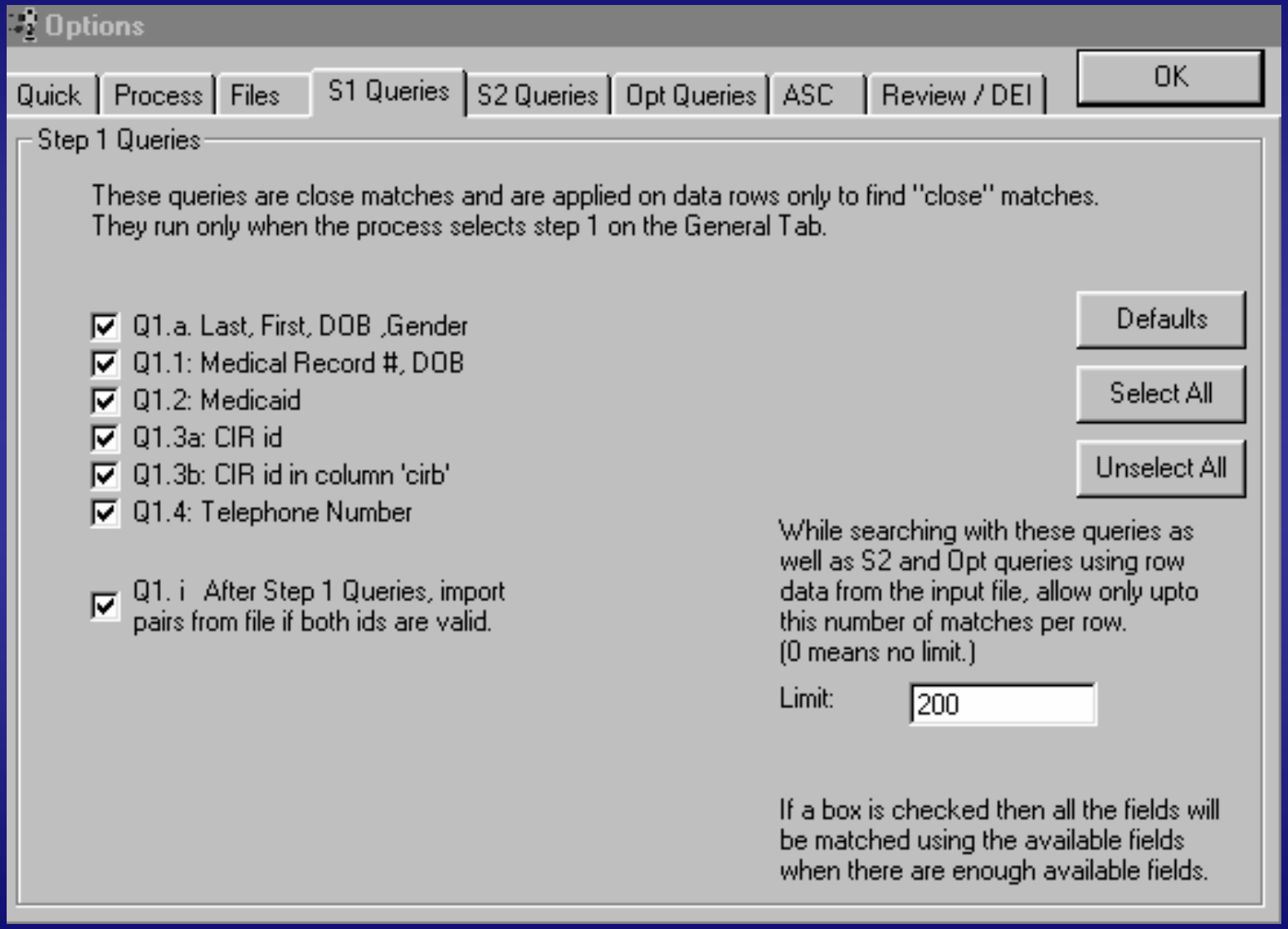
Process


Options
Quick
Process
Files
S1 Queries
S2 Queries
Opt Queries
ASC
Review / DEI
OK
General
These Options control the main flow of the process. You can select different query sets to run on the data rows from the file, and then on the row results. If you haven't selected a query set here, then it doesn't run. If you have selected a query set here, then it runs just the queries you have selected from the set, and the set runs in the way and order signified by this page.
First, run the following queries On The File Data to find as many matches as possible.
Run Step 1 Queries. Call any results 'Set 1'
Run Step 2 Queries. Put any results in 'Set 2'
Run Optional Queries. Put the results in 'Set 2'
Other Query Options:
Download History & Recommendations.
ICE: Update marked ids if files need it.
Force ICE : Always mark and need.
Add pairs implicit in row matches.
After File Rows have been searched, repeat the following queries to find more matches, and id pairs, starting with the matches from above.
Use Step 2 Queries in this process.
Use Optional Queries in this process.
Check for all pairs and matches in this order upto (and including) this step:
1. Use just the results in 'Set 1' to find more results. Put them in 'Set 2'
2. Next use results from 'Set 2' to find more results. Call them 'Set 3'.
3. Next use the results you just got, and look for their matches. ('Set 4')
4. Last use the results from that and look for matches one more time. ('Set 5')

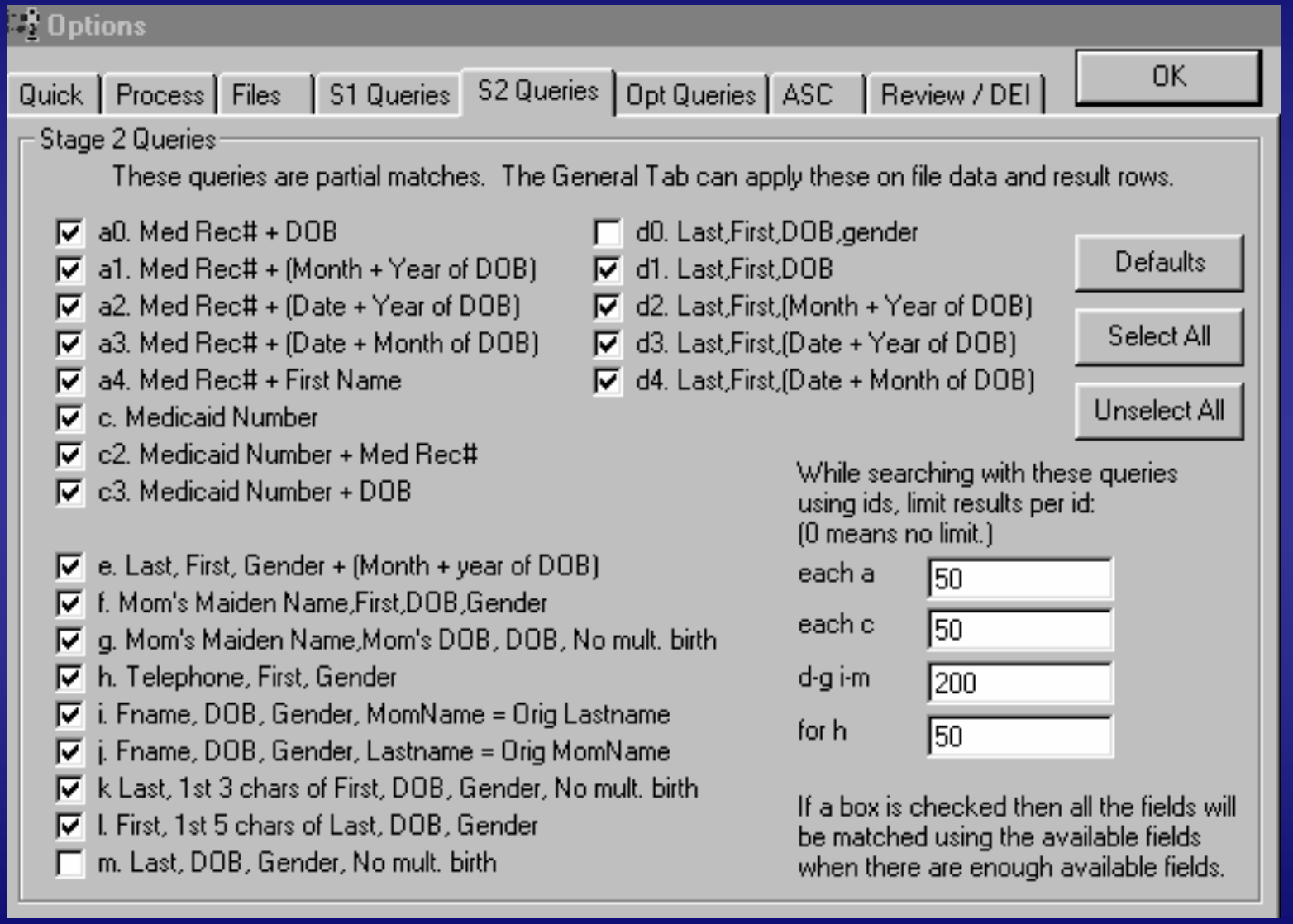
Files


Options
Quick
Process
Files
S1 Queries
S2 Queries
Opt Queries
ASC
Review / DEI
OK
Files
These options control what files you get, and the details of those files.
Generate -mas.txt File, with records found by step 1 only. (very close matches.)
Generate -pai.txt File, which contains unique possible-match ID pairs.
Add every pair that is indirectly connected by ids returned in the search.
Generate -rec.txt File, which contains all the records found relating to each input row.
Generate -not.txt File, filled with the rows from the input file which have zero matches.
Generate -med.txt File, which follows the MEDD file specification.
Generate -asc.txt File with the options on the ASC Fields tab.
Generate -recall.txt File, which follows the Recall file specification.
Generate -rrr.txt File, which follows the Recall file specification.
Generate DEI Files, which follow the DEI file specification.
Output only Recall and DEI files when reviewing records.
Include Rows from the input file in the -rec.txt file in the following cases:
For rows with zero matches
For rows with matches
Place "(X)MATCHES.FOR " before the row.
Default File Options
SQL log: (choose one)
No SQL log.
Normal SQL log.
Detailed log with results. (a big file)
Generate -log.txt File, which is the Process Log.
Ora Trace (select only when instructed to)

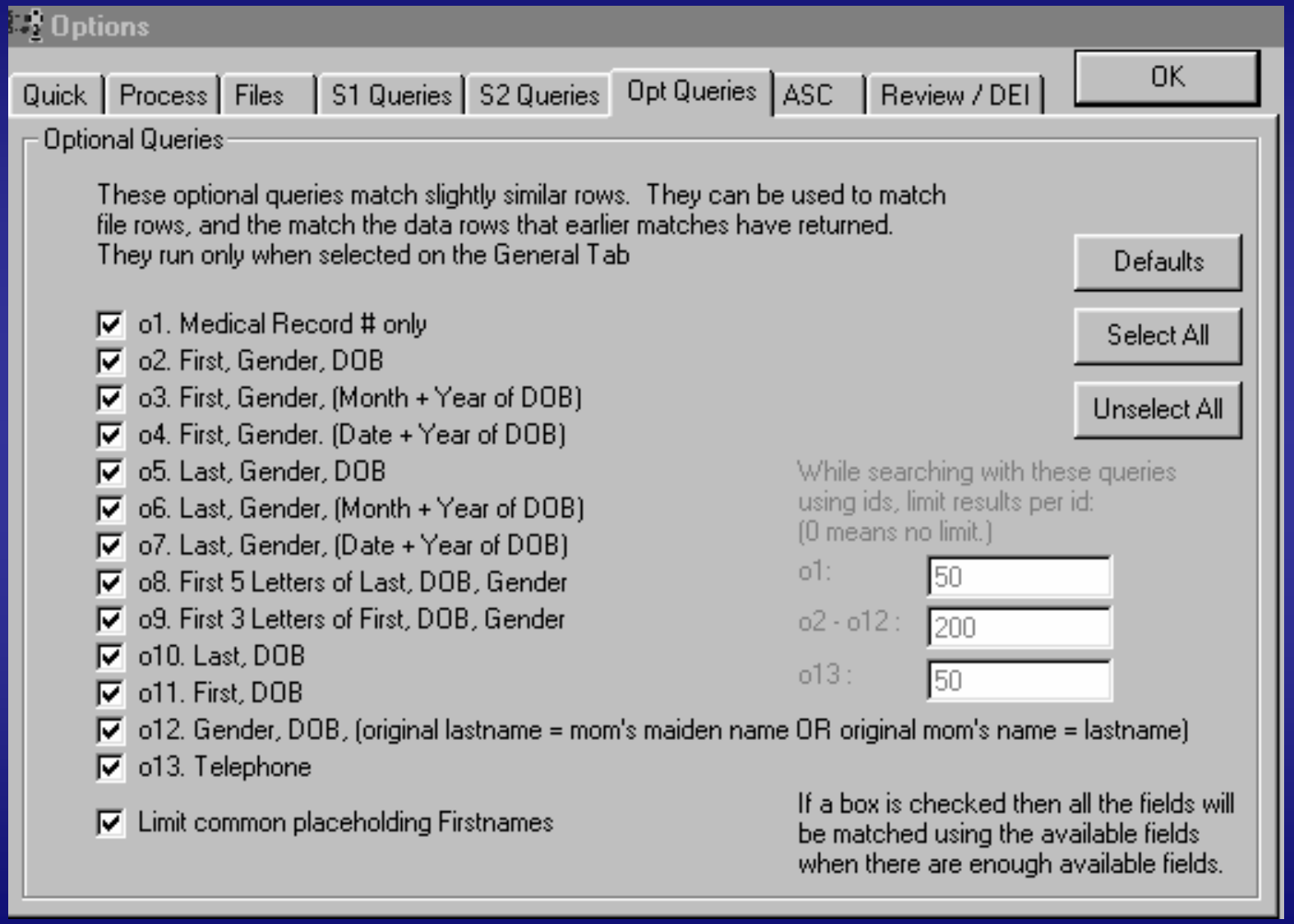
Step 1 Queries



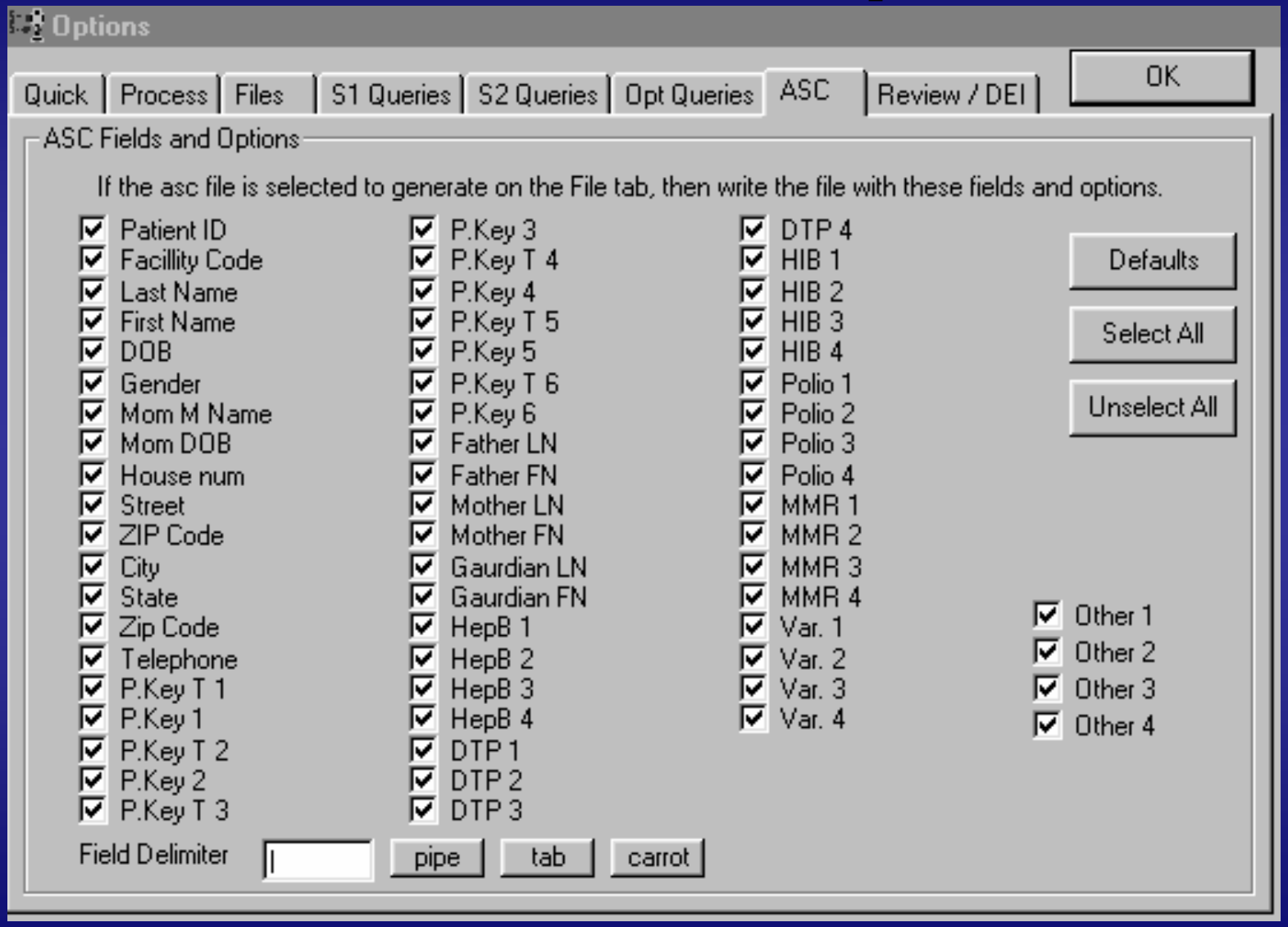
Step 2 Queries



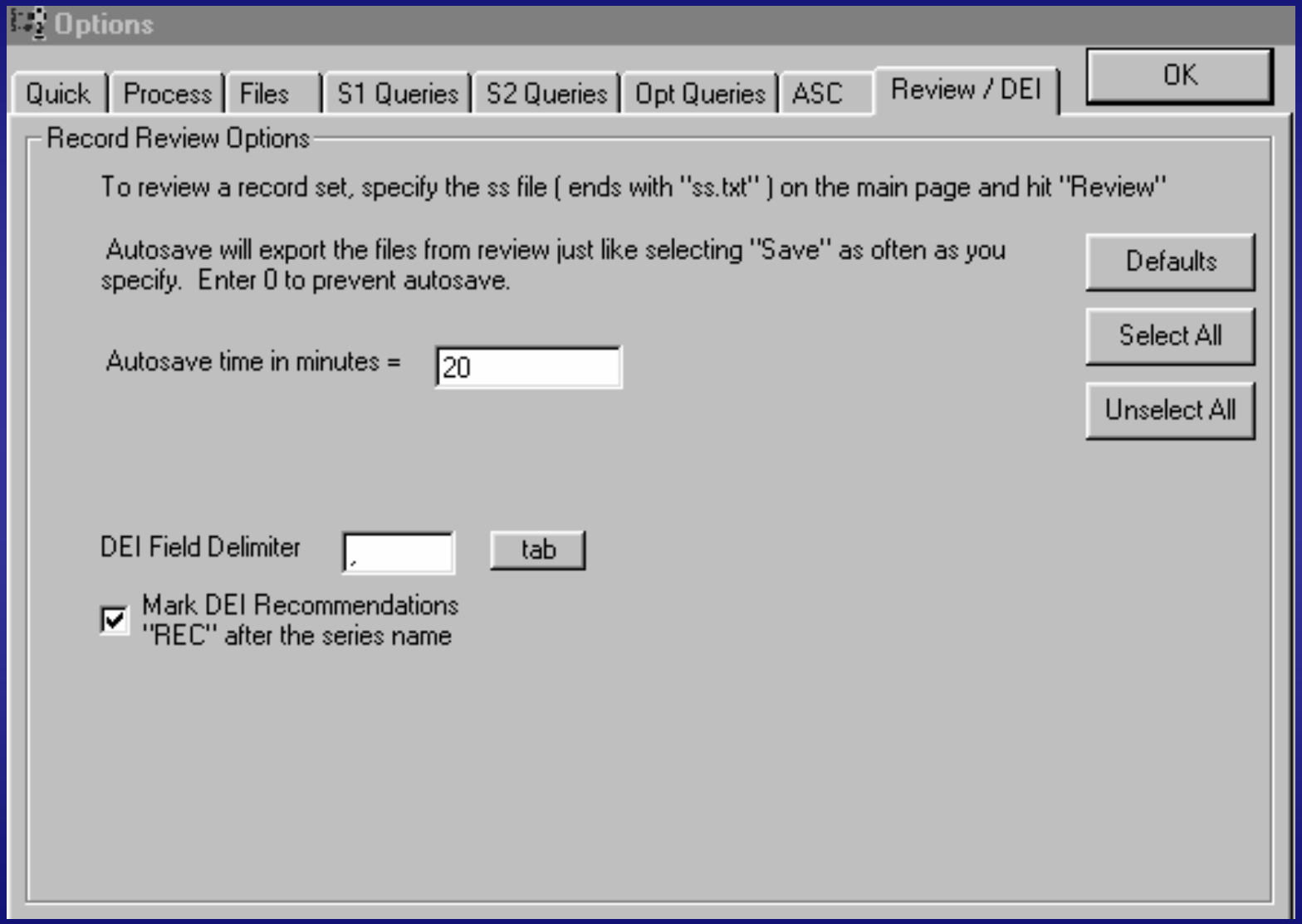
Optional Queries



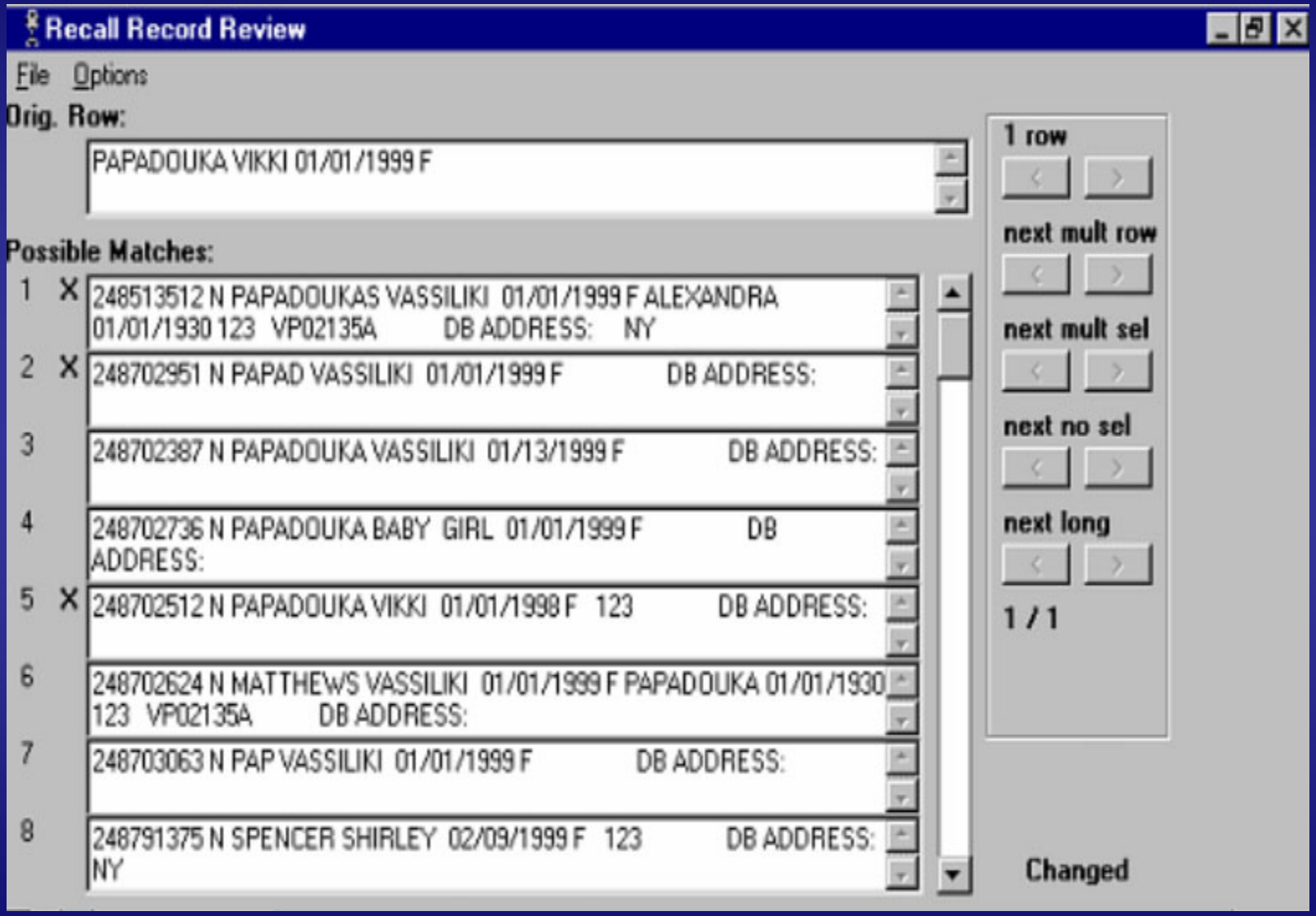
ASCI Fields & Options



Review


Options
Quick
Process
Files
S1 Queries
S2 Queries
Opt Queries
ASC
Review / DEI
OK
Record Review Options
To review a record set, specify the ss file (ends with "ss.txt") on the main page and hit "Review"
Autosave will export the files from review just like selecting "Save" as often as you specify. Enter 0 to prevent autosave.
Defaults
Select All
Unselect All
Autosave time in minutes =
20
DEI Field Delimiter
,
tab
Mark DEI Recommendations "REC" after the series name

Record Review



Conclusion

- **Smart Search is a powerful query tool which is a vital component of our de-duplication and data quality efforts, and our efforts to utilize Immunization Registry data**