

MICHAEL STACKHOUSE BIOGRAPHY

Michael Stackhouse holds a bachelor's degree from Arcadia University where he studied Business Administration, Economics, and Statistics. He is currently a MIDS student at UC Berkeley School of Information, studying Data Science. Mr. Stackhouse has extensive CDISC experience, working with both SDTM and ADaM standards as well as serving as a subject matter expert for Define.xml. Mr. Stackhouse's favorite career projects include macro and utility development, as well as molding the minds of young programmer Padawans into the programming Jedis of the future. Mr. Stackhouse is currently a Manager of Statistical Programming at Covance and father to three young, furry, four-legged children.



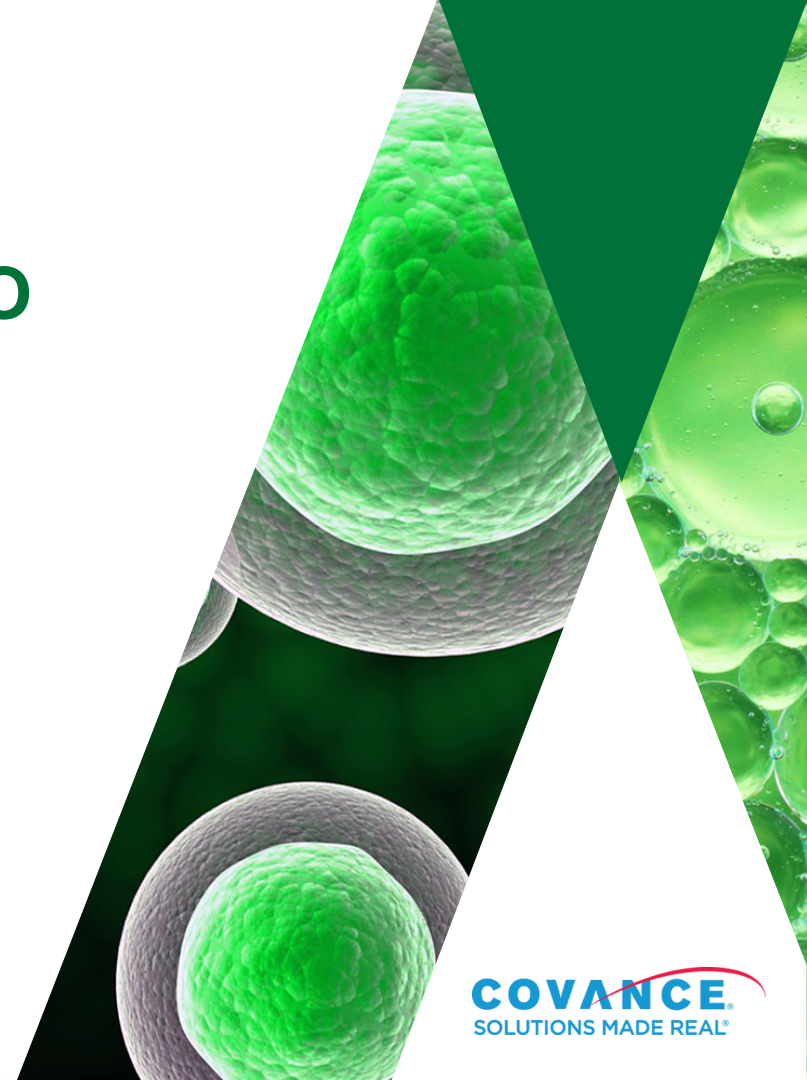
WHY SAS PROGRAMMERS SHOULD LEARN PYTHON TOO

Michael Stackhouse, Covance, Inc.



Copyright © 2018 Covance. All Rights Reserved

Public

A large, stylized graphic on the right side of the slide. It features a large green triangle pointing downwards, which is partially filled with a pattern of green and white spheres, resembling a molecular structure or a cluster of cells. The background of the slide is white, and the graphic is positioned on the right side, extending from the top to the bottom.

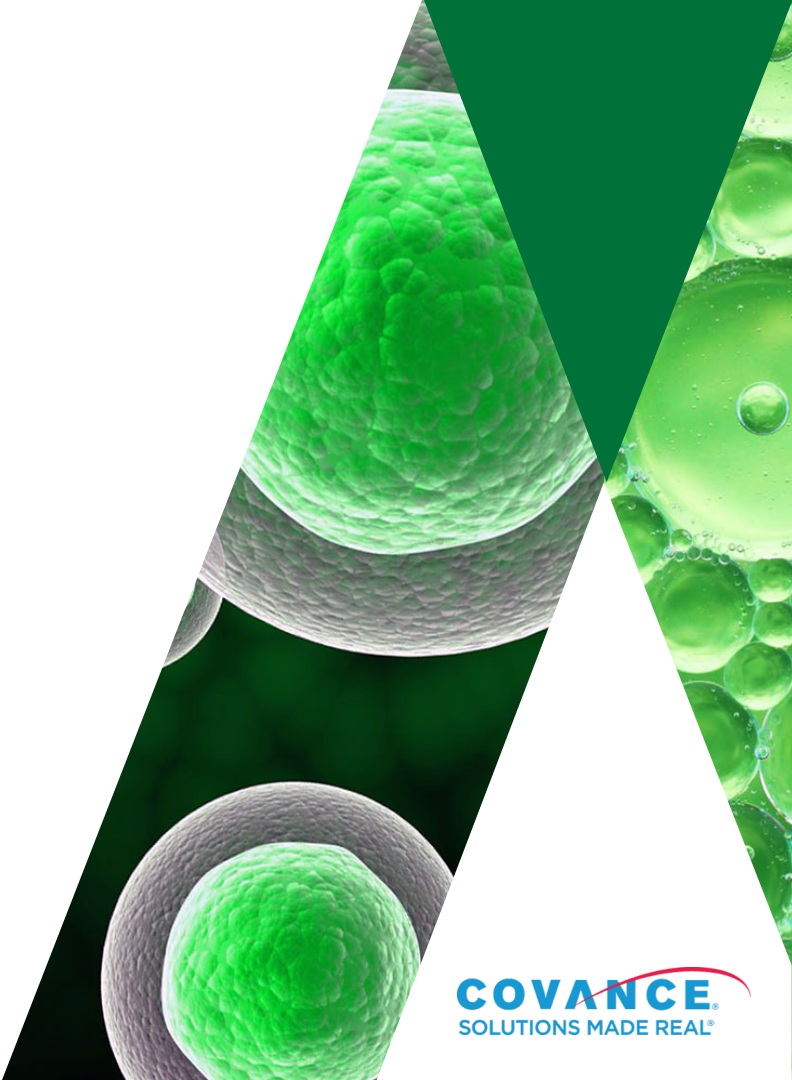
COVANCE
SOLUTIONS MADE REAL®

INTRODUCTION



Copyright © 2018 Covance. All Rights Reserved

Public

A large, stylized graphic on the right side of the slide. It features a large green triangle pointing downwards, which is partially filled with a pattern of green and white spheres, resembling a molecular structure or a cluster of cells. The background of the slide is white, and the overall theme is scientific or pharmaceutical.

COVANCE
SOLUTIONS MADE REAL®



What does Python Have to Offer?

► It's easy:

- Spoken Word Syntax is intuitive and easy to learn
- SAS is a great language for data processing but can be complex when doing tasks better suited for a general purpose programming language

► It's Powerful

- Base package has tools to handle complex tasks
- Very useful built in datatypes
- A world of 3rd party packages to increase tool base
- Can be used with Linux/Unix command line to make tasks easier



The Possibilities are endless

- ▶ A few examples of how Python can make your day to day tasks both simpler and more efficient.
 - Enhance the Command Line
 - Easily gather information
 - Update numerous files at one time

With a little bit of background and a little bit of code, you can make a huge impact on your team's work!

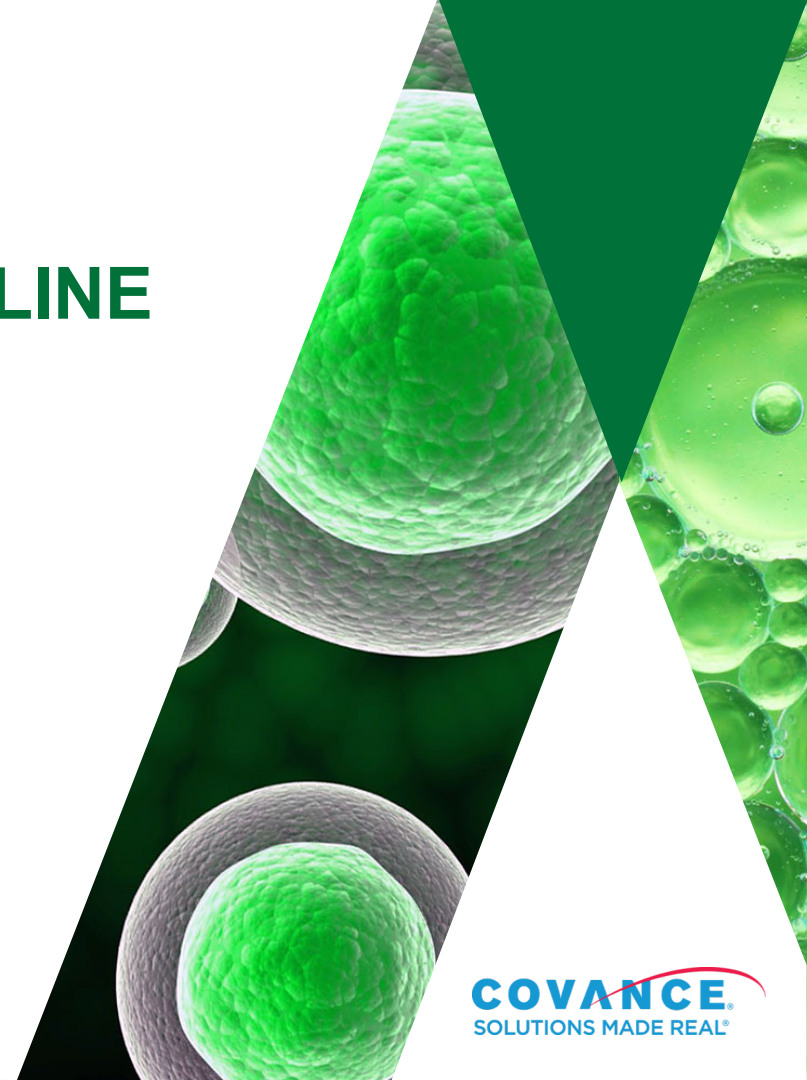


ENHANCING THE COMMAND LINE



Copyright © 2018 Covance. All Rights Reserved

Public

A large, abstract graphic on the right side of the slide. It features a dark green triangle pointing upwards, which is partially filled with a lighter green, textured, spherical shape. Below this, there are several smaller, translucent green spheres of varying sizes, some with internal textures, set against a dark green background. The overall effect is a modern, scientific, or technological aesthetic.

COVANCE
SOLUTIONS MADE REAL®

Talk Back

- ▶ Instead of memorizing command line syntax, prompt the user for input
- ▶ May make users who aren't comfortable with Linux more likely to use the tools

```
>>> my_var = raw_input("Hello! What's your name? ")  
Hello! What's your name? Mike  
>>> print my_var  
Mike
```



smart_search.py

What files do you want to check? >>> test.sas

```
#####  
# List of options #  
# i -> Ignore case #  
# v -> Search instead for NON-matching lines #  
# c -> Return instead a match of a count of matching #  
# lines for each input file #  
# l -> Print file names with a match instead #  
# of matching lines #  
# n -> Add the line number of matching lines to output #  
#####
```

String to search for >>> **Inputs**

List all options desired (hit enter for none) >>> i

Search for substring? (Y/N) >>> n

Command executed: fgrep -H -i 'Inputs' test.sas

```
test.sas:Inputs : sdtm.dm, sdtm.suppdm, sdtm.fa, sdtm.suppfa, sdtm.vs,  
sdtm.lb,
```

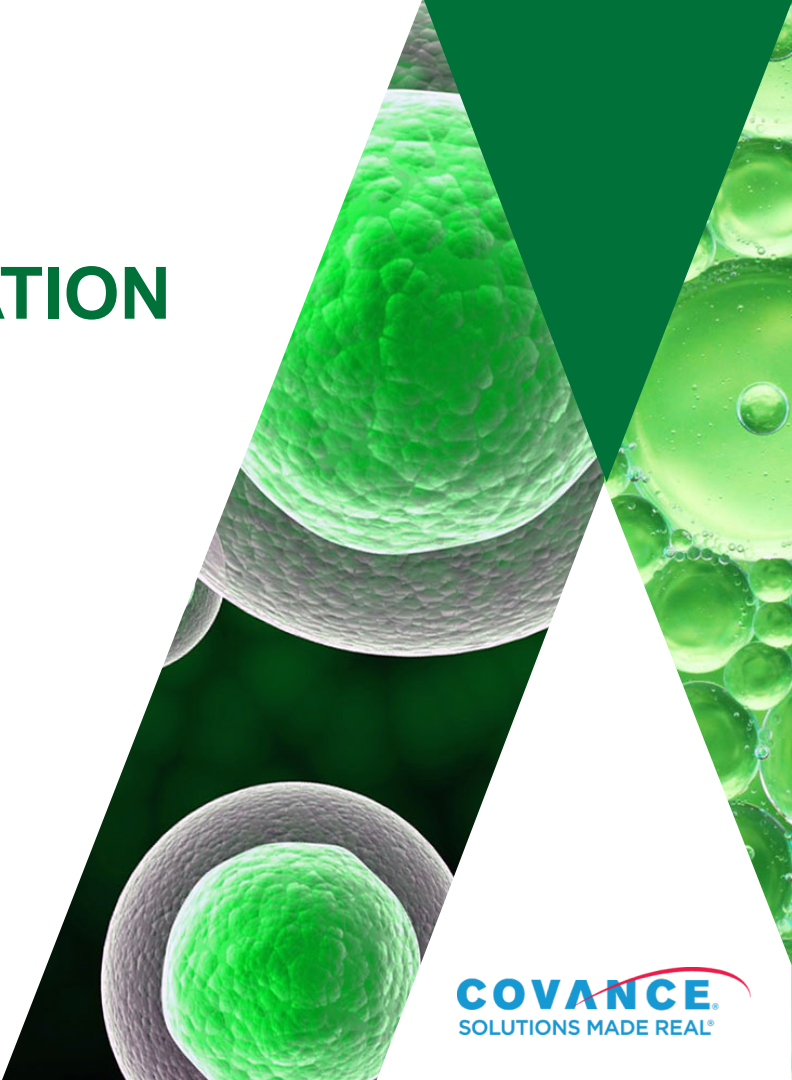


EASILY GATHERING INFORMATION



Copyright © 2018 Covance. All Rights Reserved

Public

A large, stylized graphic on the right side of the slide. It features a large green triangle pointing downwards, which is partially filled with a pattern of green and white bubbles. Below the triangle, there are several large, textured green spheres, some of which are also filled with a bubble pattern. The background of the right side is white.

COVANCE
SOLUTIONS MADE REAL®

Gathering Updates

check_date.py

Tables, listings, figures, ADaM, SDTM, or all files? adam

Please enter the date time of comparison. Use the format: MMM DD YYYY HH:NN
(i.e. Apr 25 2017 15:23)

Comparison Date? **Jun 23 2017 09:00**

Development code updated files:

adeff.sas Mon Jun 26 20:32:51 2017
adlb.sas Tue Jul 11 19:39:18 2017
adeg.sas Mon Jun 26 18:10:55 2017

QC code updated files:

gadeff.sas Fri Jun 23 20:09:03 2017
gadlb.sas Wed Jul 12 17:04:50 2017
gadmo.sas Wed Jul 12 17:33:26 2017
gadeq.sas Fri Jun 23 17:05:55 2017



Looking for Issues

ascii_check.py

Non-ascii Character Check

List the files you want to scan, separated by spaces

You can also use Linux wildcards like '?' and '*'

What files do you want to check? *.*

Print to terminal? Answering N will create a CSV file to review

(Y or N) >>> Y

File Name	Issue Found
test.txt	Non-ascii character detected at Line: 1 Column: 0.
test.txt	Non-ascii character detected at Line: 1 Column: 1.
test.txt	Non-ascii character detected at Line: 7 Column: 0.
test.txt	Non-ascii character detected at Line: 7 Column: 1.
test.txt	Non-ascii character detected at Line: 17 Column: 0.
test.txt	Non-ascii character detected at Line: 17 Column: 1.
test.xml	Non-ascii character detected at Line: 1 Column: 0.
test.xml	Non-ascii character detected at Line: 1 Column: 1.
test.xml	Non-ascii character detected at Line: 7 Column: 0.
test.xml	Non-ascii character detected at Line: 7 Column: 1.
test.xml	Non-ascii character detected at Line: 17 Column: 0.
test.xml	Non-ascii character detected at Line: 17 Column: 1.

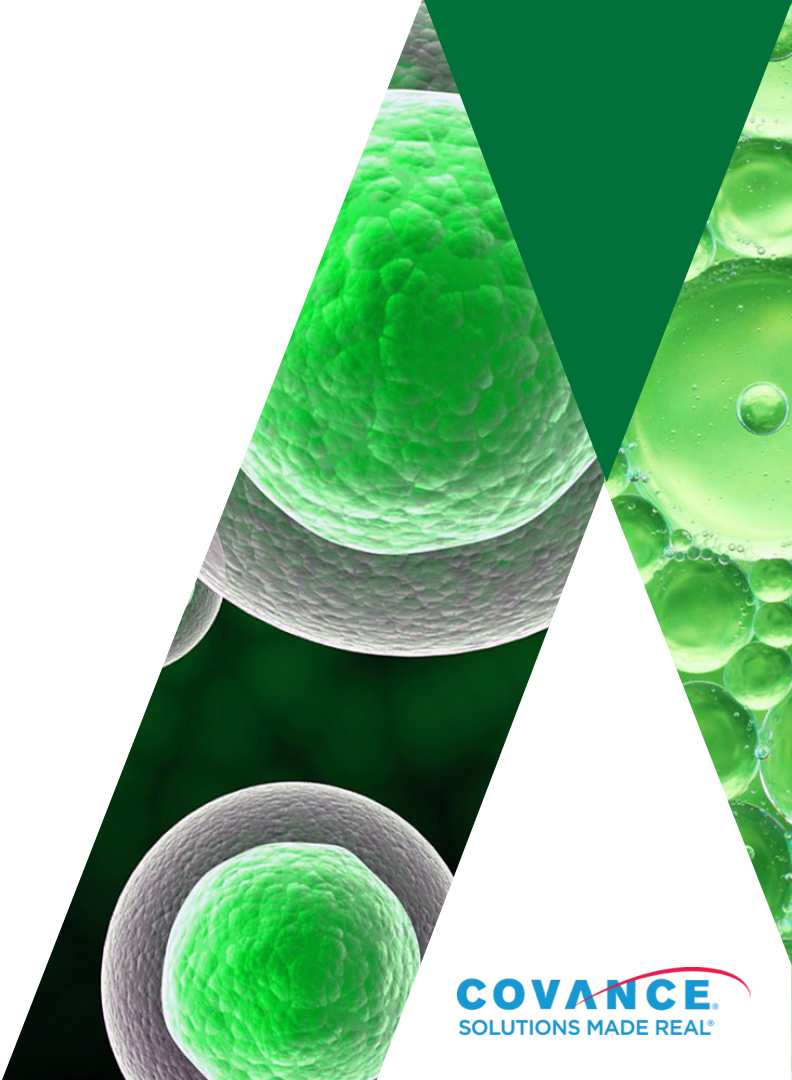


UPDATING NUMEROUS FILES AT ONCE



Copyright © 2018 Covance. All Rights Reserved

Public

A decorative graphic on the right side of the slide. It features a large green triangle pointing downwards, which contains several green, textured spheres of varying sizes. To the right of the triangle, there are green bubbles of various sizes, some with highlights, set against a white background.

COVANCE
SOLUTIONS MADE REAL®

Easily Find and Replace

- Replace Text String:

```
>>> "Hello! My name is Mike.".replace("Hello","Goodbye")  
'Goodbye! My name is Mike.'
```

- List file within directory and subdirectory:

```
os.walk(<directory>, topdown=False)
```



Easily Find and Replace

```
data x;  
  set y;  
  studyid = "MIRACLE-DRUG";  
run;
```

From text? >>> **MIRACLE-DRUG**
To text? >>> **MIRA-DRUG**

```
data x;  
  set y;  
  studyid = "MIRA-DRUG";  
run;
```



Put It All Together

► Challenge

- On your system, development and validation programs reside in the same folder with different names. On the client system, they are stored in separate folders with identical names.
- The client system requires an %INCLUDE line at the top of the file. Your system does not, so you comment it out.
- Some macros and macro variables must be renamed between systems.
- Some LIBNAMEs must be renamed between systems.



Put It All Together

► Solution

- Establish what package needs to be created (i.e. SDTM, ADaM, Tables, Listings, Figures)
- Create subdirectories to hold the program files for editing.
- Copy the programs into the subdirectories, separating development and validation in the process.
- Rename all validation programs to match their development side counterparts.
- Execute global replaces on all of the program files to uncomment the %INCLUDE line, rename the necessary macros, macro variables, and LIBNAMEs.
- Create zip files for easy delivery to the client system.

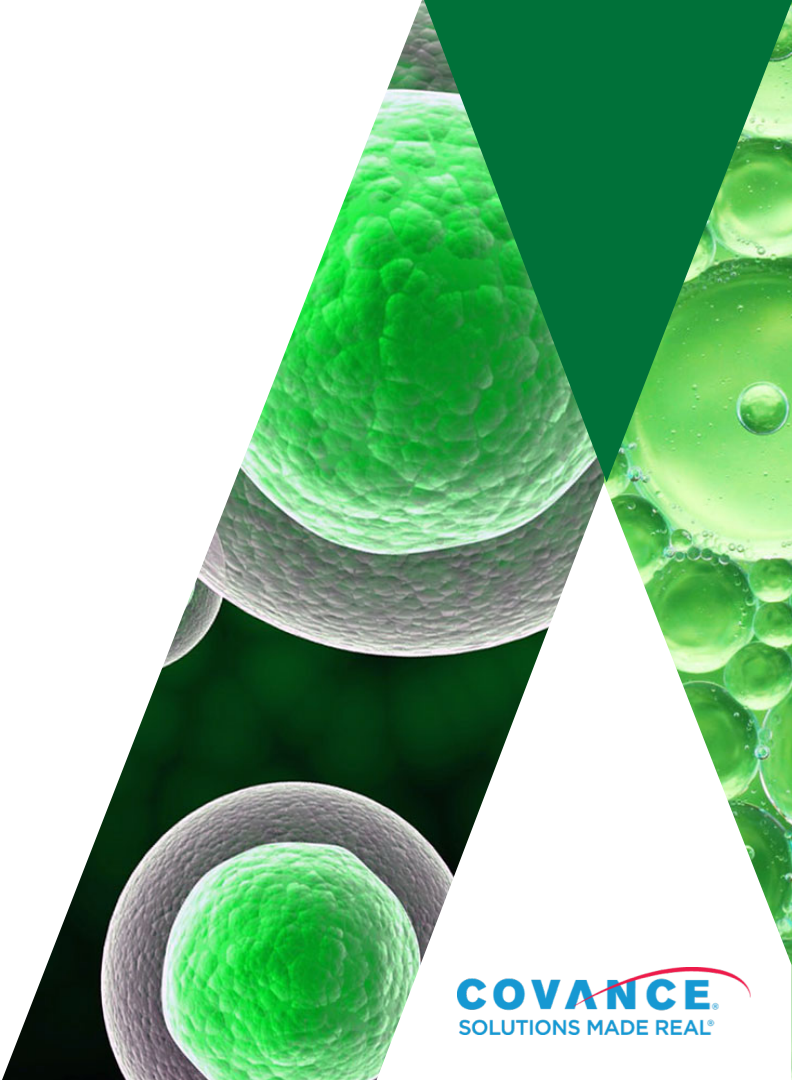


CONCLUSION AND Q&A



Copyright © 2018 Covance. All Rights Reserved

Public

A large, stylized graphic on the right side of the slide. It features a large green triangle pointing downwards, which is partially filled with a close-up image of green, textured spheres (possibly cells or bubbles) against a dark green background. The triangle is set against a white background that also features a close-up image of green bubbles.

COVANCE
SOLUTIONS MADE REAL®

Conclusion and Q&A

- ▶ Python can wrap Linux/Unix command line in a friendly UI
 - Invites less advanced users to use power of Linux/Unix command
- ▶ With little code it can gather info across directories and files
- ▶ Python can crawl across directories and make redundant updates a snap!
- ▶ It's easy to learn and with the basics you can do big things!



Conclusion and Q&A

► Questions?



Name: Michael Stackhouse
Organization: Covance, Inc.
Address: 4000 Centergreen Way # 300
City, State ZIP: Cary, NC 27513
E-mail: Michael.Stackhouse@Chiltern.com
Web: <https://www.linkedin.com/in/michael-s-stackhouse/>



Covance Inc., headquartered in Princeton, NJ, USA, is the drug development business of Laboratory Corporation of America Holdings (LabCorp). COVANCE is a registered trademark and the marketing name for Covance Inc. and its subsidiaries around the world.

