Supplement (tables, legend to supplemental figure, data on PCR)

Supplemental table S1. Demographics of patients with available *lytA* rtPCR from sputum and nasopharyngeal swab

Characteristic	n
Age, mean ± SD, years	36.7 ± 9.6
Female gender	138/222 (62.2%)
Smoker (current)	26/219 (11.9%)
New HIV diagnosis	103/222 (46.4%)
Receiving antiretroviral therapy on admission	24/222 (10.8%)
CD4 cell count, mean ± SD, cells/mm^3	131.7 ± 153.4
Receiving cotrimoxazole on admission	25/222 (11.3%)
CURB-65	
0	53 (24.0%)
1	87 (39.4%)
2	61 (27.6%)
3	18 (8.1%)
4	2 (0.9%)
Hospital death	23/213 (10.8%)

Supplemental table S2. Performance of sputum *lytA* rtPCR >10000/ml against the composite diagnostic standard

	Any sputum quality	Good sputum quality	Suboptimal sputum quality
Sensitivity	94/126 (74.6%)	25/32 (78.1%)	69/94 (73.4%)
Specificity	80/96 (83.3%)	8/10 (80.0%)	72/86 (83.7%)
PPV	94/110 (85.5%)	25/27 (92.6%)	69/83 (83.1%)
NPV	80/112 (71.4%)	8/15 (53.3%)	72/97 (74.2%)
Additional yield	16/96 (16.7%)	2/10 (20.0%)	14/86 (16.2%)

The composite diagnostic standard was considered positive if any of blood culture, induced sputum culture or Gram stain, urine antigen, whole blood *lytA* rtPCR revealed pneumococcus or if NPS *lytA* rtPCR was >8000 copies/ml.

PPV – positive predictive value; NPV – negative predictive value; additional yield=1-specificity (i.e. patients with sputum *lytA* rtPCR >10000/ml and a negative composite diagnostic standard, represents “false-positives”)

Supplemental table S3. Comparison between genomic load measured by *lytA* rtPCR from sputum and nasopharyngeal swab (NPS) in HIV-uninfected patients

	Sputum *lytA* rtPCR>10000/ml (any sputum quality) (n=35)	Sputum *lytA* rtPCR>10000/ml (good sputum quality) (n=10)	Sputum *lytA* rtPCR>10000/ml (suboptimal sputum quality) (n=25)	NPS *lytA* rtPCR>8000/ml (n=35)
Sensitivity	11/14 (78.6%)	5/6 (83.3%)	6/8 (75.0%)	9/14 (64.3%)
Specificity	12/21 (57.1%)	4/4 (100%)	8/17 (47.1%)	17/21 (81.0%)
PPV	11/20 (55.0%)	5/5 (100%)	6/15 (40.0%)	9/13 (69.2%)
NPV	12/15 (80.0%)	4/5 (80%)	8/10 (80.0%)	17/22 (77.3%)
Additional yield	9/21 (42.9%)	0 (0%)	9/17 (52.9%)	4/21 (19.0%)

Composite diagnostic standard: any of positive blood culture, induced sputum culture or Gram stain, urine antigen, whole blood *lytA* rtPCR; for this comparison criterion NPS *lytA* rtPCR >8000 copies/ml was not included in the composite diagnostic

PPV – positive predictive value; NPV – negative predictive value; additional yield=1-specificity (i.e. patients with sputum *lytA* rtPCR >10000/ml or NPS sputum *lytA* rtPCR >8000/ml, respectively, and a negative composite diagnostic standard, represents “false-positives”

Supplemental figure legend

Supplemental figure. Receiver-operating characteristic (ROC) curves for pneumococcal etiology

Pneumococcal etiology based on presence of composite diagnostic, i.e. any of positive blood culture, induced sputum culture or Gram stain, urine antigen, whole blood *lytA* rtPCR. ROC curves for nasopharyngeal (NP) swab, good quality sputum, suboptimal quality sputum and any quality sputum

PCR from whole blood

The PCR analyses from whole blood were performed at the Fondation Mérieux- Emerging Pathogens Laboratory, Lyon, France. Nucleic acids were extracted in a BSL-2 laboratory and PCR analysis were performed in a BSL-1 laboratory. Samples had been frozen at -70°C prior to shipment, shipment was on dried ice and storage at -80°C prior to processing. PCR mixes were prepared in a clean room under a PCR cabinet.

Nucleic acids extraction: nucleic acids were extracted from 200 µL of whole blood using the QIAamp® DNA Blood Mini Kit (Qiagen, Netherlands) and eluted in 100 µL of elution buffer according to the manufacturer's instructions.

Real-time PCR for the detection of *S. pneumoniae/H. influenzaeB/S. aureus*: 10 µl of extracted nucleic acids are added to the iQ Multiplex Powermix (Bio-Rad, USA) containing 270nM of each primer and probe, and submitted to 95° C for 10 min then 8sec at 95°C then 34 sec at 60°C for 40 cycles. The fluorophores used for each specific probes are FAM for *S. aureus*, Cy5 for *S. pneumoniae* and HEX for *H. influenzae* B.

PCR from sputum and nasopharyngeal (NP) swab

The methods described in the online supplement to Albrich et al. CID 2012; 54(5):601-9, were applied to the sputum and NP swab. Sputum was homogenized by addition of 1ml of Sputagest and was vortexed afterwards. Samples were then frozen at -70°C prior to shipment, shipment was on dried ice and then specimens were kept frozen at -80 ºC prior to use in real-time PCR (rtPCR) studies. rtPCR from sputum and NP swabs were performed at Wyeth Vaccine Research, Pearl River, NY, USA, between 2 and 4 years after specimen collection.

Both specimens were analyzed using a duplex quantitative rtPCR assay designed to detect the *lytA* and *ply* genes simultaneously.

The nucleic acid was extracted from both specimen types as follows: For sputum samples, 100 µl of sputum samples was first pre-treated with 0.15% DTT at 37 ºC, followed by Proteinase K (PK) treatment. After digestion, DNA was isolated and purified using ABI PRISMTM 6700 Automated Nucleic Workstation as per manufacturer's instructions. For NP swab specimens, 100 µl of sample was directly digested by proteinase K and DNA isolated and purified as described previously.

Rt PCR assays with primer-probe sets targeting the *lytA* and *ply* genes (see Table below) were performed under the following conditions: 10 µl DNA was mixed with 10 µl of 2X Fast PCR mix (Applied Biosystems) containing 250 nM of primers and probes from both lytA and ply genes. The PCR assay was performed using a 7900 Fast Real-Time PCR System with Fast 96-well Block Module (Applied Biosystems) using these cycling parameters: 10 min at 95 ºC, followed by 45 cycles of 15 s at 95 ºC and 1 min at 60 ºC. Each assay plate also included negative control wells with no template and wells with positive control (*S.pneumoniae* serotype 1) DNA.

The endpoint of the assay was the cycle threshold (Ct) or the cycle at which DNA amplification crossed the pre-determined baseline. Following the completion of the runs, all data was analysed using the SDS version v2.3 application software (Applied Biosystems). All reactions with Ct > 45 were designated as negative because this number is near the assay's limit of detection. Samples with Ct <45 for *lytA* were considered as an *S. pneumoniae* positive sample. The semi-quantitation of DNA copies/ml was estimated based on the Ct differences observed between samples to known amount positive DNA (*S. pneumoniae* serotype 1) control and assuming a reaction efficiency of 2.

Table: Real-time PCR primers and probes

Gene Primer/Probe	**Name**	**Sequence 5'→3'**
lytA primer	AU-306 FP	ACGCAATCTAGCAGATGAAGCA
	AU-379 RP	CGTGCGTTTTAATTCCAGCTAA
ply primer	PN-893 FP	TGCAGAGCGTCCTTTGGTCTAT
	PN-973 RP	CTCTTACTCGTGGTTTCCAACTTGA
lytA probe	AU-355	TCCCTGTATCAAGCGTTTTCGGCAAA
ply probe	PN-917	TGGCGCCCATAAGCAACACTCGAA